A Mathematical Introduction to Signals and Systems

Time and frequency domain representations of signals

Andrew D. Lewis

This version: 2016/11/26

Preface for series

The subject of signals and systems, particularly linear systems, is by now an entrenched part of the curriculum in many engineering disciplines, particularly electrical engineering. Furthermore, the offshoots of signals and systems theory—e.g., control theory, signal processing, and communications theory—are themselves well-developed and equally basic to many engineering disciplines. As many a student will agree, the subject of signals and systems is one with a reliance on tools from many areas of mathematics. However, much of this mathematics is not revealed to undergraduates, and necessarily so. Indeed, a complete accounting of what is involved in signals and systems theory would take one, at times quite deeply, into the fields of linear algebra (and to a lesser extent, algebra in general), real and complex analysis, measure and probability theory, and functional analysis. Indeed, in signals and systems theory, many of these topics are woven together in surprising and often spectacular ways. The existing texts on signals and systems theory, and there is a true abundance of them, all share the virtue of presenting the material in such a way that it is comprehensible with the bare minimum background.

Should I bother reading these volumes?

This virtue comes at a cost, as it must, and the reader must decide whether this cost is worth paying. Let us consider a concrete example of this, so that the reader can get an idea of the sorts of matters the volumes in this text are intended to wrestle with. Consider the function of time

$$f(t) = \begin{cases} \mathrm{e}^{-t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

In the text (Example 13.1.3–2) we shall show that, were one to represent this function in the frequency domain with frequency represented by ν, we would get

$$\hat{f}(\nu) = \int_{\mathbb{R}} f(t)\mathrm{e}^{-2\mathrm{i}\pi\nu t}\,\mathrm{d}t = \frac{1}{1 + 2\mathrm{i}\pi\nu}.$$

The idea, as discussed in Chapter 9, is that $\hat{f}(\nu)$ gives a representation of the "amount" of the signal present at the frequency ν. Now, it is desirable to be able to reconstruct f from $\hat{f}$, and we shall see in Section 13.2 that this is done via the formula

$$f(t) \text{``=''} \int_{\mathbb{R}} \hat{f}(\nu)\mathrm{e}^{2\mathrm{i}\pi\nu t}\,\mathrm{d}\nu. \tag{FT}$$

The easiest way to do the integral is, of course, using a symbolic manipulation program. I just tried this with *Mathematica*®, and I was told it could not do the computation. Indeed, the integral *does not converge*! Nonetheless, in many tables of Fourier transforms (that is what the preceding computations are about), we are told that the integral in (FT) does indeed produce $f(t)$. Are the tables wrong? Well, no.

But they are only correct when one understands exactly what the right-hand side of (FT) means. What it means is that the integral converges, *in* $\mathsf{L}^2(\mathbb{R};\mathbb{C})$ to f. Let us say some things about the story behind this that are of a general nature, and apply to many ideas in signal and system theory, and indeed to applied mathematics as a whole.

1. The story, it is the story of the L^2-Fourier transform, is not completely trivial. It requires *some* delving into functional analysis at least, and some background in integration theory, if one wishes to understand that "L" stands for "Lebesgue," as in "Lebesgue integration." At its most simple-minded level, the theory is certainly understandable by many undergraduates. Also, at its most simple-minded level, it raises more questions than it answers.
2. The story, even at the most simple-minded level alluded to above, takes some time to deliver. The full story takes *a lot* of time to deliver.
3. It is not necessary to fully understand the story, perhaps even the most simple-minded version of it, to be a user of the technology that results.
4. By understanding the story well, one is led to new ideas, otherwise completely hidden, that are practically useful. In control theory, quadratic regulator theory, and in signal processing, the Kalman filter, are examples of this. ***missing stuff*** has written about this, in fact.
5. The full story of the L^2-Fourier transform, and the issues stemming from it, directly or otherwise, are beautiful.

The nature of the points above, as they relate to this series, are as follows. Points 1 and 2 indicate why the story cannot be told to all undergraduates, or even most graduate students. Point 3 indicates why it is okay that the story not be told to everyone. Point 4 indicates why it is important that the story be told to someone. Point 5 should be thought of as a sort of benchmark as to whether the reader should bother with understanding what is in this series. Here is how to apply it. If one reads the assertion that this is a beautiful story, and their reaction is, "Okay, but there better be a payoff," or, "So what?" or, "Beautiful to who?" then perhaps they should steer clear of this series. If they read the assertion that this is a beautiful story, and respond with, "Really? Tell me more," then I hope they enjoy these books. They were written for such readers. Of course, most readers' reactions will fall somewhere in between the above extremes. Such readers will have to sort out for themselves whether the volumes in this series lie on the right side, for them, of being worth reading. For these readers I will say that this series is *heavily* biased towards readers who react in an unreservedly positive manner to the assertions of intrinsic beauty.

For readers skeptical of assertions of the usefulness of mathematics, an interesting pair of articles concerning this is [**RWH:80**].***missing stuff***

What is the best way of getting through this material?

Now that a reader has decided to go through with understanding what is in these volumes, they are confronted with actually doing so: a possibly nontrivial matter, depending on their starting point. Let us break down our advice according to the background of the reader.

I look at the tables of contents, and very little seems familiar. Clearly if nothing seems familiar at all, then a reader should not bother reading on until they have acquired an at least passing familiarity with some of the topics in the book. This can be done by obtaining an undergraduate degree in electrical engineering (or similar), or pure or applied mathematics.

If a reader already possess an undergraduate degree in mathematics or engineering, then certainly some of the following topics will appear to be familiar: linear algebra, differential equations, some transform analysis, Fourier series, system theory, real and/or complex analysis. However, it is possible that they have not been taught in a manner that is sufficiently broad or deep to quickly penetrate the texts in this series. That is to say, relatively inexperienced readers will find they have some work to do, even to get into topics with which they have some familiarity. The best way to proceed in these cases depends, to some extent, on the nature of one's background.

I am familiar with some or all of the applied topics, but not with the mathematics. For readers with an engineering background, even at the graduate level, the depth with which topics are covered in these books is perhaps a little daunting. The best approach for such readers is to select the applied topic they wish to learn more about, and then use the text as a guide. When a new topic is initiated, it is clearly stated what parts of the book the reader is expected to be familiar with. The reader with a more applied background will find that they will not be able to get far without having to unravel the mathematical background almost to the beginning. Indeed, readers with a typical applied background will typically be lacking a good background in linear algebra and real analysis. Therefore, they will need to invest a good deal of effort acquiring some quite basic background. At this time, they will quickly be able to ascertain whether it is worth proceeding with reading the books in this series.

I am familiar with some or all of the mathematics, but not with the applied topics. Readers with an undergraduate degree in mathematics will fall into this camp, and probably also some readers with a graduate education in engineering, depending on their discipline. They may want to skim the relevant background material, just to see what they know and what they don't know, and then proceed directly to the applied topics of interest.

I am familiar with most of the contents. For these readers, the series is one of reference books.

Comments on organisation

In the current practise of teaching areas of science and engineering connected with mathematics, there is much emphasis on "just in time" delivery of mathematical ideas and techniques. Certainly I have employed this idea myself in the classroom, without thinking much about it, and so apparently I think it a good thing. However, the merits of the "just in time" approach in written work are, in my opinion, debatable. The most glaring difficulty is that the same mathematical ideas can be "just in time" for multiple non-mathematical topics. This can even happen in a single one semester course. For example—to stick to something germane to this series—are differential equations "just in time" for general system theory? for modelling? for feedback control theory? The answer is, "For all of them," of course. However, were one to choose one of these topics for a "just in time" written delivery of the material, the presentation would immediately become awkward, especially in the case where that topic were one that an instructor did not wish to cover in class.

Another drawback to a "just in time" approach in written work is that, when combined with the corresponding approach in the classroom, a connection, perhaps unsuitably strong, is drawn between an area of mathematics and an area of application of mathematics. Given that one of the strengths of mathematics is to facilitate the connecting of seemingly disparate topics, inside and outside of mathematics proper, this is perhaps an overly simplifying way of delivering mathematical material. In the "just simple enough, but not too simple" spectrum, we fall on the side of "not too simple."

For these reasons and others, the material in this series is generally organised according to its mathematical structure. That is to say, mathematical topics are treated independently and thoroughly, reflecting the fact that they have life independent of any specific area of application. We do not, however, slavishly follow the Bourbaki[1] ideals of logical structure. That is to say, we do allow ourselves the occasional forward reference when convenient. However, we are certainly careful to maintain the standards of deductive logic that currently pervade the subject of "mainstream" mathematics. We also do not slavishly follow the Bourbaki dictum of starting with the most general ideas, and proceeding to the more specific. While there is something to be said for this, we feel that for the subject and intended readership of this series, such an approach would be unnecessarily off-putting.

[1] Bourbaki refers to "Nicolas Bourbaki," a pseudonym given (by themselves) to a group of French mathematicians who, beginning in mid-1930's, undertook to rewrite the subject of mathematics. Their dictums include presenting material in a completely logical order, where no concepts is referred to before being defined, and starting developments from the most general, and proceeding to the more specific. The original members include Henri Cartan, André Weil, Jean Delsarte, Jean Dieudonné, and Claude Chevalley, and the group later counted such mathematicians as Roger Godement, Jean-Pierre Serre, Laurent Schwartz, Emile Borel, and Alexander Grothendieck among its members. They have produced eight books on fundamental subjects of mathematics.

This version: 2016/11/26

Table of Contents

1 Set theory and terminology **1**

1.1 Sets . . . 3

1.1.1 Definitions and examples . . . 3

1.1.2 Unions and intersections . . . 5

1.1.3 Finite Cartesian products . . . 7

1.2 Relations . . . 10

1.2.1 Definitions . . . 10

1.2.2 Equivalence relations . . . 12

1.3 Maps . . . 14

1.3.1 Definitions and notation . . . 14

1.3.2 Properties of maps . . . 16

1.3.3 Graphs and commutative diagrams . . . 19

1.4 Indexed families of sets and general Cartesian products . . . 25

1.4.1 Indexed families and multisets . . . 25

1.4.2 General Cartesian products . . . 27

1.4.3 Sequences . . . 28

1.4.4 Directed sets and nets . . . 28

1.5 Some words about proving things . . . 30

1.5.1 Legitimate proof techniques . . . 30

1.5.2 Improper proof techniques . . . 31

2 Real numbers and their properties **34**

2.1 Construction of the real numbers . . . 36

2.1.1 Construction of the rational numbers . . . 36

2.1.2 Construction of the real numbers from the rational numbers 41

2.2 Properties of the set of real numbers . . . 46

2.2.1 Algebraic properties of $\mathbb{R}$ . . . 46

2.2.2 The total order on $\mathbb{R}$ . . . 50

2.2.3 The absolute value function on $\mathbb{R}$ . . . 53

2.2.4 Properties of $\mathbb{Q}$ as a subset of $\mathbb{R}$ . . . 54

2.2.5 The extended real line . . . 58

2.2.6 sup and inf . . . 60

2.2.7 Notes . . . 61

2.3 Sequences in $\mathbb{R}$ . . . 63

2.3.1 Definitions and properties of sequences . . . 63

2.3.2 Some properties equivalent to the completeness of $\mathbb{R}$ . . . 65

2.3.3 Tests for convergence of sequences . . . 68

2.3.4 lim sup and lim inf . . . 69

2.3.5 Multiple sequences 72
2.3.6 Algebraic operations on sequences 74
2.3.7 Convergence using $\mathbb{R}$-nets 75
2.3.8 A first glimpse of Landau symbols 80
2.3.9 Notes 82
2.4 Series in $\mathbb{R}$ 84
2.4.1 Definitions and properties of series 84
2.4.2 Tests for convergence of series 90
2.4.3 e and π 94
2.4.4 Doubly infinite series 98
2.4.5 Multiple series 100
2.4.6 Algebraic operations on series 101
2.4.7 Series with arbitrary index sets 104
2.4.8 Notes 106
2.5 Subsets of $\mathbb{R}$ 110
2.5.1 Open sets, closed sets, and intervals 110
2.5.2 Partitions of intervals 114
2.5.3 Interior, closure, boundary, and related notions 115
2.5.4 Compactness 121
2.5.5 Connectedness 126
2.5.6 Sets of measure zero 126
2.5.7 Cantor sets 130
2.5.8 Notes 132

3 Functions of a real variable **134**
3.1 Continuous $\mathbb{R}$-valued functions on $\mathbb{R}$ 137
3.1.1 Definition and properties of continuous functions 137
3.1.2 Discontinuous functions 141
3.1.3 Continuity and operations on functions 145
3.1.4 Continuity, and compactness and connectedness 147
3.1.5 Monotonic functions and continuity 150
3.1.6 Convex functions and continuity 153
3.1.7 Piecewise continuous functions 159
3.2 Differentiable $\mathbb{R}$-valued functions on $\mathbb{R}$ 163
3.2.1 Definition of the derivative 163
3.2.2 The derivative and continuity 167
3.2.3 The derivative and operations on functions 170
3.2.4 The derivative and function behaviour 175
3.2.5 Monotonic functions and differentiability 183
3.2.6 Convex functions and differentiability 190
3.2.7 Piecewise differentiable functions 196
3.2.8 Notes 197
3.3 $\mathbb{R}$-valued functions of bounded variation 199
3.3.1 Functions of bounded variation on compact intervals 199
3.3.2 Functions of bounded variation on general intervals 203

3.3.3 Bounded variation and operations on functions 206
3.3.4 Saltus functions 210
3.3.5 The saltus function for a function of locally bounded variation 213
3.4 The Riemann integral 217
3.4.1 Step functions 217
3.4.2 The Riemann integral on compact intervals 219
3.4.3 Characterisations of Riemann integrable functions on compact intervals 221
3.4.4 The Riemann integral on noncompact intervals 228
3.4.5 The Riemann integral and operations on functions 234
3.4.6 The Fundamental Theorem of Calculus and the Mean Value Theorems 239
3.4.7 The Cauchy principal value 245
3.4.8 Notes 247
3.5 Sequences and series of ℝ-valued functions 248
3.5.1 Pointwise convergent sequences 248
3.5.2 Uniformly convergent sequences 249
3.5.3 Dominated and bounded convergent sequences 252
3.5.4 Series of ℝ-valued functions 254
3.5.5 Some results on uniform convergence of series 255
3.5.6 The Weierstrass Approximation Theorem 257
3.5.7 Swapping limits with other operations 263
3.5.8 Notes 266
3.6 Some ℝ-valued functions of interest 267
3.6.1 The exponential function 267
3.6.2 The natural logarithmic function 269
3.6.3 Power functions and general logarithmic functions 271
3.6.4 Trigonometric functions 275
3.6.5 Hyperbolic trigonometric functions 282

4 Algebraic structures 285
4.1 Groups 287
4.1.1 Definitions and basic properties 287
4.1.2 Subgroups 290
4.1.3 Quotients 292
4.1.4 Group homomorphisms 294
4.1.5 The symmetric group 296
4.2 Fields 303
4.2.1 Definitions and basic properties 303
4.2.2 Fraction fields 304
4.2.3 Subfields, field homomorphisms, and characteristic 305
4.3 Vector spaces 309
4.3.1 Definitions and basic properties 309
4.3.2 Subspaces 312
4.3.3 Linear independence 315

4.3.4 Basis and dimension . . . 320
4.3.5 Intersections, sums, and products . . . 327
4.3.6 Complements and quotients . . . 334
4.3.7 Complexification of $\mathbb{R}$-vector spaces . . . 337
4.3.8 Extending the scalars for a vector space . . . 340
4.3.9 Notes . . . 342

5 Measure theory and integration 344
5.1 Some motivation for abstract measure theory and integration . . . 347
5.1.1 The Jordan measure and its limitations . . . 347
5.1.2 Some limitations of the Riemann integral . . . 356
5.1.3 An heuristic introduction to the Lebesgue integral . . . 359
5.1.4 Notes . . . 361
5.2 Measurable sets . . . 362
5.2.1 Algebras and σ-algebras . . . 362
5.2.2 Algebras and σ-algebras generated by families of subsets . . . 364
5.2.3 Products of measurable spaces . . . 369
5.3 Measures . . . 374
5.3.1 Functions on families of subsets . . . 374
5.3.2 Outer measures, measures, and their relationship . . . 379
5.3.3 Complete measures and completions of measures . . . 387
5.3.4 Outer and inner measures associated to a measure . . . 389
5.3.5 Probability measures . . . 396
5.3.6 Product measures . . . 398
5.3.7 Signed measures . . . 404
5.3.8 Complex measures . . . 411
5.3.9 Vector measures . . . 414
5.3.10 Spaces of positive, signed, complex, and vector measures . . . 416
5.3.11 Notes . . . 418
5.4 Lebesgue measure on $\mathbb{R}$ . . . 419
5.4.1 The Lebesgue outer measure and the Lebesgue measure on $\mathbb{R}$ 419
5.4.2 Borel sets in $\mathbb{R}$ as examples of Lebesgue measurable sets . . . 423
5.4.3 Further properties of the Lebesgue measure on $\mathbb{R}$ . . . 428
5.4.4 Notes . . . 435
5.5 Lebesgue measure on $\mathbb{R}^n$ . . . 436
5.5.1 The Lebesgue outer measure and the Lebesgue measure on $\mathbb{R}^n$ 436
5.5.2 Borel sets in $\mathbb{R}^n$ as examples of Lebesgue measurable sets . . . 438
5.5.3 Further properties of the Lebesgue measure on $\mathbb{R}^n$ . . . 442
5.5.4 Lebesgue measure on $\mathbb{R}^n$ as a product measure . . . 449
5.5.5 Coverings of subsets of $\mathbb{R}^n$ . . . 451
5.5.6 The Banach–Tarski Paradox . . . 455
5.5.7 Notes . . . 466
5.6 Measurable functions . . . 468
5.6.1 General measurable maps and functions . . . 468
5.6.2 Measurability and operations on functions . . . 474

5.6.3 Sequences of measurable functions . . . 479
5.6.4 ℂ- and vector-valued measurable functions . . . 486
5.6.5 Simple functions and approximations of measurable functions 490
5.6.6 Topological characterisations of convergence for sequences of measurable functions . . . 493
5.7 Integration on measure spaces . . . 500
5.7.1 Definition of the integral . . . 500
5.7.2 The integral and operations on functions . . . 510
5.7.3 Limit theorems . . . 514
5.7.4 Integration with respect to probability measures . . . 520
5.7.5 Topological characterisations of limit theorems . . . 523
5.7.6 Image measure and integration by image measure . . . 525
5.7.7 The integral for ℂ- and vector-valued functions . . . 527
5.7.8 Integration with respect to signed, complex, and vector measures . . . 531
5.7.9 Notes . . . 542
5.8 Integration on products . . . 543
5.8.1 The product measure by integration . . . 543
5.8.2 The integral on product spaces . . . 546
5.8.3 Fubini's Theorem . . . 547
5.9 The single-variable Lebesgue integral . . . 556
5.9.1 Lebesgue measurable functions . . . 556
5.9.2 The (conditional) Lebesgue integral . . . 560
5.9.3 Properties of the Lebesgue integral . . . 563
5.9.4 Swapping operations with the Lebesgue integral . . . 569
5.9.5 Locally Lebesgue integrable functions . . . 572
5.9.6 Absolute continuity . . . 574
5.9.7 The Fundamental Theorem of Calculus for the Lebesgue integral . . . 581
5.9.8 Lebesgue points . . . 593
5.9.9 Maximal functions . . . 595
5.9.10 The change of variables formula . . . 595
5.9.11 Topological characterisations of the deficiencies of the Riemann integral . . . 598
5.10 Notes . . . 601

6 Banach spaces **602**
6.1 Definitions and properties of normed vector spaces . . . 604
6.1.1 Norms and seminorms . . . 604
6.1.2 Open and closed subsets of normed vector spaces . . . 610
6.1.3 Subspaces, direct sums, and quotients . . . 614
6.2 Sequences in normed vector spaces . . . 619
6.2.1 Definitions and properties of sequences . . . 619
6.2.2 Algebraic operations on sequences . . . 621
6.2.3 Multiple sequences . . . 621

6.3 Completeness and completions . . . 623
6.3.1 Completeness (Banach spaces) . . . 623
6.3.2 Why completeness is important . . . 627
6.3.3 Completeness and direct sums and quotients . . . 629
6.3.4 Completions . . . 630
6.4 Series in normed vector spaces . . . 634
6.4.1 Definitions and properties of series . . . 634
6.4.2 Absolute and unconditional convergence . . . 636
6.4.3 Algebraic operations on series . . . 645
6.4.4 Multiple series . . . 647
6.4.5 Cesàro convergence of sequences and series . . . 648
6.4.6 Series in normed vector spaces with arbitrary index sets . . . 650
6.4.7 Notes . . . 655
6.5 Continuous maps between normed vector spaces . . . 656
6.5.1 General continuous maps between normed vector spaces . . . 656
6.5.2 Continuous linear maps between normed vector spaces . . . 659
6.5.3 Induced topologies on continuous linear maps . . . 664
6.5.4 The Open Mapping Theorem and Closed Graph Theorem . . . 670
6.6 Topology of normed vector spaces . . . 672
6.6.1 Properties of balls in normed vector spaces . . . 672
6.6.2 Interior, closure, boundary, etc. . . . 673
6.6.3 Compactness . . . 678
6.6.4 Closed subspaces . . . 680
6.6.5 Bases for normed vector spaces . . . 689
6.6.6 Notes . . . 692
6.7 Examples of Banach spaces . . . 693
6.7.1 The p-norms on $\mathbb{F}^n$ . . . 693
6.7.2 Banach spaces of sequences . . . 700
6.7.3 Banach spaces of direct sums . . . 709
6.7.4 Banach spaces of continuous functions on $\mathbb{R}$ . . . 713
6.7.5 Banach spaces of continuous functions on metric spaces . . . 720
6.7.6 Banach spaces of continuous functions on locally compact topological spaces . . . 721
6.7.7 Banach spaces of integrable functions on $\mathbb{R}$ . . . 721
6.7.8 Banach spaces of integrable functions on measure spaces . . . 732
6.7.9 Banach spaces of measures . . . 732
6.7.10 Notes . . . 734

7 Hilbert spaces **736**
7.1 Definitions and properties of inner product spaces . . . 738
7.1.1 Inner products and semi-inner products . . . 738
7.1.2 Inner product spaces as normed vector spaces . . . 741
7.1.3 Orthogonality . . . 746
7.1.4 Hilbert spaces and their subspaces . . . 749
7.1.5 Minimising distance to a set . . . 754

7.1.6 Norms . . . 757
7.2 Continuous maps between inner product spaces . . . 760
7.2.1 The dual of an inner product space . . . 760
7.2.2 Particular aspects of continuity for inner product spaces . . . 763
7.2.3 Notes . . . 764
7.3 Orthonormal bases in Hilbert spaces . . . 765
7.3.1 General definitions and results . . . 765
7.3.2 Finite orthonormal sets and finite Hilbert bases . . . 776
7.3.3 Countable orthonormal sets and countable Hilbert bases . . . 780
7.3.4 Generalised Fourier series . . . 786
7.3.5 Classification of Hilbert spaces . . . 789
7.3.6 Notes . . . 793

8 Signals in the time-domain 795
8.1 Time, signals, and elementary properties . . . 797
8.1.1 Examples and problems involving time-domain signals . . . 797
8.1.2 Time . . . 800
8.1.3 Time-domain signals: basic definitions . . . 803
8.1.4 Elementary transformations of signals . . . 806
8.1.5 Causal and acausal signals . . . 812
8.1.6 Periodic and harmonic signals . . . 813
8.1.7 Characterising time-domain signals . . . 817
8.1.8 Notes . . . 819
8.2 Spaces of discrete-time signals . . . 821
8.2.1 The vector space structure of general discrete-time signals . . . 821
8.2.2 Discrete-time signal spaces characterised by their values . . . 822
8.2.3 The ℓ^p-spaces . . . 823
8.2.4 Periodic discrete-time signal spaces . . . 824
8.2.5 Other characteristics of discrete-time signals . . . 825
8.2.6 Inclusions of discrete-time signal spaces . . . 826
8.3 Spaces of continuous-time signals . . . 830
8.3.1 The vector space structure of general continuous-time signals 830
8.3.2 Spaces of continuous signals . . . 832
8.3.3 The L^p-spaces . . . 833
8.3.4 Periodic continuous-time signal spaces . . . 836
8.3.5 Other characteristics of continuous-time signals . . . 838
8.3.6 Inclusions of continuous-time signal spaces . . . 841
8.3.7 Some other useful classes of continuous-time signals . . . 845
8.3.8 Notes . . . 847

9 Signals in the frequency-domain 851
9.1 Frequency and substance . . . 853
9.1.1 The electromagnetic spectrum . . . 853
9.1.2 Emission and absorption spectra . . . 855
9.1.3 Determining the composition of a distant star . . . 856

9.2 Sound and music . . . 858
9.2.1 Sound identification using frequency . . . 858
9.2.2 A little music theory . . . 859
9.2.3 Notes . . . 863
9.3 Signal transmission schemes . . . 864
9.3.1 A general communication problem . . . 864
9.3.2 Amplitude modulation . . . 865
9.3.3 Frequency modulation . . . 867
9.3.4 Phase modulation . . . 868
9.4 System identification using frequency response . . . 870
9.4.1 System modelling . . . 870
9.4.2 A simple example . . . 871
9.4.3 A not-so-simple example . . . 872
9.4.4 Notes . . . 873
9.5 Frequency-domains and signals in the frequency-domain . . . 874
9.5.1 Frequency . . . 874
9.5.2 Frequency-domain signals: basic definitions and properties . 875
9.5.3 Spaces of discrete frequency-domain signals . . . 878
9.5.4 Spaces of continuous frequency-domain signals . . . 879
9.6 An heuristic introduction to the four Fourier transforms . . . 880
9.6.1 Periodic continuous-time signals . . . 880
9.6.2 Aperiodic continuous-time signals . . . 882
9.6.3 Periodic discrete-time signals . . . 885
9.6.4 Aperiodic discrete-time signals . . . 887
9.6.5 Notes . . . 888

10 Distributions in the time-domain **890**
10.1 Motivation for distributions . . . 893
10.1.1 The impulse . . . 893
10.1.2 Differentiating nondifferentiable signals . . . 895
10.1.3 What should a good theory of distributions achieve? . . . 896
10.1.4 Some caveats about what is written about the delta-signal . . 897
10.1.5 Notes . . . 897
10.2 Distributions . . . 899
10.2.1 Test signals . . . 899
10.2.2 Definition of distributions . . . 902
10.2.3 Locally integrable signals are distributions . . . 903
10.2.4 The support and singular support of a distribution . . . 905
10.2.5 Convergence of distributions . . . 907
10.2.6 Differentiation of distributions . . . 910
10.2.7 Integration of distributions . . . 915
10.2.8 Distributions depending on parameters . . . 917
10.2.9 Fubini's Theorem for distributions . . . 921
10.2.10 Some deeper properties of distributions . . . 923
10.2.11 The order of a distribution . . . 926

10.2.12 Measures as distributions 929
10.3 Tempered distributions 932
10.3.1 The Schwartz space of test signals 932
10.3.2 Definition of tempered distributions 935
10.3.3 Properties of tempered distributions 941
10.3.4 Tempered distributions depending on parameters 943
10.3.5 Fubini's Theorem for tempered distributions 946
10.3.6 Some deeper properties of tempered distributions 948
10.4 Integrable distributions 953
10.4.1 Bounded test signals 953
10.4.2 Definition of integrable distributions 954
10.4.3 Properties of integrable distributions 957
10.4.4 Some deeper properties of integrable distributions 958
10.4.5 Notes 964
10.5 Distributions with compact support 965
10.5.1 The set of infinitely differentiable test signals 965
10.5.2 Definition of distributions with compact support 966
10.5.3 Properties of distributions with compact support 970
10.5.4 Distributions with compact support depending on parameters972
10.5.5 Fubini's Theorem for distributions with compact support . . 974
10.5.6 Some deeper properties of distributions with compact support975
10.5.7 Some constructions with delta-signals 977
10.6 Ultradistributions 984
10.6.1 The space $\mathscr{Z}(\mathbb{R};\mathbb{F})$ of test signals 984
10.6.2 Definition of ultradistributions 986
10.6.3 Properties of ultradistributions 987
10.6.4 Some deeper properties of ultradistributions 988
10.7 Periodic distributions 989
10.7.1 Periodic test signals 989
10.7.2 Definition of periodic distributions 992
10.7.3 Properties of periodic distributions 995
10.7.4 Some deeper properties of periodic distributions 997
10.8 Inclusions between signals, test signals, and generalised signals . . . 1000

11 Convolution **1004**
11.1 Convolution of signals: Definitions, basic properties, and examples . 1006
11.1.1 Convolution for aperiodic continuous-time signals 1006
11.1.2 Convolution for periodic continuous-time signals 1016
11.1.3 Convolution for aperiodic discrete-time signals 1028
11.1.4 Convolution for periodic discrete-time signals 1033
11.1.5 Notes 1038
11.2 Convolvable pairs of signals 1039
11.2.1 Convolution in $\mathsf{L}^1(\mathbb{R};\mathbb{F})$ 1039
11.2.2 Convolution between $\mathsf{L}^p(\mathbb{R};\mathbb{F})$ and $\mathsf{L}^q(\mathbb{R};\mathbb{F})$ 1053
11.2.3 Convolution of signals with restrictions on their support . . . 1055

11.2.4 Convolution in $\mathsf{L}^1_{\mathrm{per},T}(\mathbb{R};\mathbb{F})$ . . . 1070
11.2.5 Convolution in $\mathsf{L}^p_{\mathrm{per},T}(\mathbb{R};\mathbb{F})$ . . . 1072
11.2.6 Convolution in $\ell^1(\mathbb{Z}(\Delta);\mathbb{F})$ . . . 1073
11.2.7 Convolution in $\ell^p(\mathbb{Z}(\Delta);\mathbb{F})$ . . . 1075
11.2.8 Convolution and regularity . . . 1076
11.2.9 Notes . . . 1079
11.3 Approximation and regularisation . . . 1081
11.3.1 Approximate identities on $\mathbb{R}$ . . . 1081
11.3.2 Approximate identities on $\mathbb{R}_{\geq 0}$ . . . 1095
11.3.3 Periodic approximate identities . . . 1098
11.3.4 Regularisation of signals on $\mathbb{R}$ . . . 1108
11.3.5 Regularisation of periodic signals . . . 1111
11.3.6 Regularisation of generalised signals . . . 1111

12 The continuous-discrete Fourier transform **1113**
12.1 The L^1-CDFT . . . 1115
12.1.1 Definitions and computations . . . 1115
12.1.2 Properties of the CDFT . . . 1122
12.1.3 Differentiation, integration, and the CDFT . . . 1127
12.1.4 Decay of the CDFT . . . 1129
12.1.5 Convolution, multiplication, and the L^1-CDFT . . . 1134
12.2 Inversion of the CDFT . . . 1139
12.2.1 Preparatory work . . . 1139
12.2.2 Fourier series . . . 1142
12.2.3 Divergence of Fourier series . . . 1147
12.2.4 Pointwise convergence of Fourier series . . . 1167
12.2.5 Uniform convergence of Fourier series . . . 1174
12.2.6 Gibbs' phenomenon . . . 1179
12.2.7 Cesàro summability . . . 1183
12.2.8 The CDFT and approximate identities . . . 1185
12.2.9 Notes . . . 1189
12.3 The L^2-CDFT . . . 1194
12.3.1 The Hilbert basis of harmonic signals . . . 1194
12.3.2 The inverse L^2-CDFT . . . 1199
12.3.3 The relationship between various notions of convergence . . 1201
12.3.4 Convolution, multiplication, and the L^2-CDFT . . . 1202
12.3.5 The Uncertainty Principle for the CDFT . . . 1203
12.3.6 Notes . . . 1203

13 The continuous-continuous Fourier transform **1207**
13.1 The L^1-CCFT . . . 1209
13.1.1 Definitions and computations . . . 1209
13.1.2 Properties of the CCFT . . . 1214
13.1.3 Differentiation, integration, and the CCFT . . . 1216

13.1.4 Decay of the CCFT 1219
13.1.5 Convolution, multiplication, and the L^1-CCFT 1221
13.1.6 Alternative formulae for the CCFT 1222
13.1.7 Notes 1223
13.2 Inversion of the CCFT 1226
13.2.1 Preparatory work 1226
13.2.2 The Fourier integral 1230
13.2.3 Divergence of Fourier integrals 1234
13.2.4 Pointwise convergence of Fourier integrals 1236
13.2.5 Uniform convergence of Fourier integrals 1242
13.2.6 Gibbs' phenomenon 1247
13.2.7 Cesàro means 1248
13.2.8 The CCFT and approximate identities 1250
13.2.9 Notes 1257
13.3 The L^2-CCFT 1259
13.3.1 Definition of the L^2-CCFT 1259
13.3.2 Properties of the L^2-CCFT 1262
13.3.3 The inverse L^2-CCFT 1264
13.3.4 Computation of the L^2-CCFT 1267
13.3.5 Convolution, multiplication, and the L^2-CCFT 1272
13.3.6 One version of the Paley–Weiner Theorem 1273
13.3.7 Notes 1280
13.4 The CCFT for tempered distributions 1283
13.4.1 The strategy for defining the CCFT of a distribution 1283
13.4.2 The Fourier transform of Schwartz test signals 1283
13.4.3 Definitions and computations 1284
13.4.4 Properties of the CCFT for tempered distributions 1285
13.4.5 Inversion of the CCFT for tempered distributions 1288
13.4.6 The CCFT for distributions with compact support 1289
13.4.7 The CCFT for periodic distributions 1292

14 Discrete-time Fourier transforms 1294
14.1 The discrete-continuous Fourier transform 1296
14.1.1 Definition of the ℓ^1-DCFT 1296
14.1.2 Properties of the DCFT 1299
14.1.3 The effect of coefficient decay on the DCFT 1301
14.1.4 Convolution, multiplication, and the DCFT 1302
14.1.5 Inversion of the DCFT 1303
14.1.6 The ℓ^2-DCFT 1304
14.1.7 The DCFT for signals of slow growth 1309
14.1.8 The DCFT for general signals 1311
14.2 The discrete-discrete Fourier transform 1313
14.2.1 The DDFT signal spaces 1313
14.2.2 Definition of the DDFT 1315
14.2.3 Properties of the DDFT 1318

14.2.4 Convolution, multiplication, and the DDFT 1320
14.2.5 Inversion of the DDFT 1322
14.2.6 The fast Fourier transform 1322
14.2.7 Notes 1327

Chapter 1

Set theory and terminology

The principle purpose of this chapter is to introduce the mathematical notation and language that will be used in the remainder of these volumes. Much of this notation is standard, or at least the notation we use is generally among a collection of standard possibilities. In this respect, the chapter is a simple one. However, we also wish to introduce the reader to some elementary, although somewhat abstract, mathematics. The secondary objective behind this has three components.

1. We aim to provide a somewhat rigorous foundation for what follows. This means being fairly clear about defining the (usually) somewhat simple concepts that arise in the chapter. Thus "intuitively clear" concepts like sets, subsets, maps, etc., are given a fairly systematic and detailed discussion. It is at least interesting to know that this can be done. And, if it is not of interest, it can be sidestepped at a first reading.
2. This chapter contains some results, and many of these require very simple proofs. We hope that these simple proofs might be useful to readers who are new to the world where everything is proved. Proofs in other chapters in these volumes may not be so useful for achieving this objective.
3. The material is standard mathematical material, and should be known by anyone purporting to love mathematics.

Do I need to read this chapter? Readers who are familiar with standard mathematical notation (e.g., who understand the symbols $\in$, $\subseteq$, $\cup$, $\cap$, $\times$, $f\colon S \to T$, $\mathbb{Z}_{>0}$, and $\mathbb{Z}$) can simply skip this chapter in its entirety. Some ideas (e.g., relations, orders, Zorn's Lemma) may need to be referred to during the course of later chapters, but this is easily done.

Readers not familiar with the above standard mathematical notation will have some work to do. They should certainly read Sections 1.1, 1.2, and 1.3 closely enough that they understand the language, notation, and main ideas. And they should read enough of Section **??** that they know what objects, familiar to them from their being human, the symbols $\mathbb{Z}_{>0}$ and $\mathbb{Z}$ refer to. The remainder of the material can be overlooked until it is needed later. •

Contents

1.1 Sets 3

1.1.1 Definitions and examples . . . 3
1.1.2 Unions and intersections . . . 5
1.1.3 Finite Cartesian products . . . 7
1.2 Relations . . . 10
1.2.1 Definitions . . . 10
1.2.2 Equivalence relations . . . 12
1.3 Maps . . . 14
1.3.1 Definitions and notation . . . 14
1.3.2 Properties of maps . . . 16
1.3.3 Graphs and commutative diagrams . . . 19
1.4 Indexed families of sets and general Cartesian products . . . 25
1.4.1 Indexed families and multisets . . . 25
1.4.2 General Cartesian products . . . 27
1.4.3 Sequences . . . 28
1.4.4 Directed sets and nets . . . 28
1.5 Some words about proving things . . . 30
1.5.1 Legitimate proof techniques . . . 30
1.5.2 Improper proof techniques . . . 31

Section 1.1

Sets

The basic ingredient in modern mathematics is the set. The idea of a set is familiar to everyone at least in the form of "a collection of objects." In this section, we shall not really give a definition of a set that goes beyond that intuitive one. Rather we shall accept this intuitive idea of a set, and move forward from there. This way of dealing with sets is called ***naïve set theory***. There are some problems with naïve set theory, as described in Section **??**, and these lead to a more formal notion of a set as an object that satisfies certain axioms, those given in Section **??**. However, these matters will not concern us much at the moment.

Do I need to read this section? Readers familiar with basic set theoretic notation can skip this section. Other readers should read it, since it contains language, notation, and ideas that are absolutely commonplace in these volumes. •

1.1.1 Definitions and examples

First let us give our working definition of a set. A ***set*** is, for us, a well-defined collection of objects. Thus one can speak of everyday things like "the set of red-haired ladies who own yellow cars." Or one can speak of mathematical things like "the set of even prime numbers." Sets are therefore defined by describing their ***members*** or ***elements***, i.e., those objects that are in the set. When we are feeling less formal, we may refer to an element of a set as a ***point*** in that set. The set with no members is the ***empty set***, and is denoted by $\emptyset$. If S is a set with member x, then we write $x \in S$. If an object x is *not* in a set S, then we write $x \notin S$.

1.1.1 Examples (Sets)

1. If S is the set of even prime numbers, then $2 \in S$.
2. If S is the set of even prime numbers greater than 3, then S is the empty set.
3. If S is the set of red-haired ladies who own yellow cars and if $x =$ Ghandi, then $x \notin S$. •

If it is possible to write the members of a set, then they are usually written between braces $\{\ \}$. For example, the set of prime numbers less that 10 is written as $\{2, 3, 5, 7\}$ and the set of physicists to have won a Fields Prize as of 2005 is {Edward Witten}.

A set S is a ***subset*** of a set T if $x \in S$ implies that $x \in T$. We shall write $S \subseteq T$, or equivalently $T \supseteq S$, in this case. If $x \in S$, then the set $\{x\} \subseteq S$ with one element, namely x, is a ***singleton***. Note that x and $\{x\}$ are different things. For example, $x \in S$ and $\{x\} \subseteq S$. If $S \subseteq T$ and if $T \subseteq S$, then the sets S and T are ***equal***, and we write $S = T$. If two sets are not equal, then we write $S \neq T$. If $S \subseteq T$ and if $S \neq T$, then S is a ***proper*** or ***strict*** subset of T, and we write $S \subset T$ if we wish to emphasise this fact.

1.1.2 Notation (Subsets and proper subsets) We adopt a particular convention for denoting subsets and proper subsets. That is, we write $S \subseteq T$ when S is a subset of T, allowing for the possibility that $S = T$. When $S \subseteq T$ and $S \neq T$ we write $S \subset T$. In this latter case, many authors will write $S \subsetneq T$. We elect not to do this. The convention we use is consistent with the convention one normally uses with inequalities. That is, one normally writes $x \leq y$ and $x < y$. It is not usual to write $x \lneq y$ in the latter case. •

Some of the following examples may not be perfectly obvious, so may require sorting through the definitions.

1.1.3 Examples (Subsets)

1. For any set S, $\emptyset \subseteq S$ (see Exercise 1.1.1).
2. $\{1, 2\} \subseteq \{1, 2, 3\}$.
3. $\{1, 2\} \subset \{1, 2, 3\}$.
4. $\{1, 2\} = \{2, 1\}$.
5. $\{1, 2\} = \{2, 1, 2, 1, 1, 2\}$. •

A common means of defining a set is to define it as the subset of an existing set that satisfies conditions. Let us be slightly precise about this. A ***one-variable predicate*** is a statement which, in order that its truth be evaluated, needs a single argument to be specified. For example, $P(x)$ = "x is blue" needs the single argument x in order that it be decided whether it is true or not. We then use the notation

$$\{x \in S \mid P(x)\}$$

to denote the members x of S for which the predicate P is true when evaluated at x. This is read as something like, "the set of x's in S such that $P(x)$ holds."

For sets S and T, the ***relative complement*** of T in S is the set

$$S - T = \{x \in S \mid x \notin T\}.$$

Note that for this to make sense, we do not require that T be a subset of S. It is a common occurrence when dealing with complements that one set be a subset of another. We use different language and notation to deal with this. If S is a set and if $T \subseteq S$, then $S \setminus T$ denotes the ***absolute complement*** of T in S, and is defined by

$$S \setminus T = \{x \in S \mid x \notin T\}.$$

Note that, if we forget that T is a subset of S, then we have $S \setminus T = S - T$. Thus $S - T$ is the more general notation. Of course, if $A \subseteq T \subseteq S$, one needs to be careful when using the words "absolute complement of A," since one must say whether one is taking the complement in T or the larger complement in S. For this reason, we prefer the notation we use rather the commonly encountered notation A^{C} or A' to refer to the absolute complement. Note that one should not talk about the absolute complement to a set, without saying within which subset the complement is being taken. To do so would imply the existence of "a set containing all sets," an object that leads one to certain paradoxes (see Section **??**).

A useful set associated with every set S is its ***power set***, by which we mean the set

$$\mathbf{2}^S = \{A \mid A \subseteq S\}.$$

The reader can investigate the origins of the peculiar notation in Exercise 1.1.3.

1.1.2 Unions and intersections

In this section we indicate how to construct new sets from existing ones.

Given two sets S and T, the ***union*** of S and T is the set $S \cup T$ whose members are members of S *or* T. The ***intersection*** of S and T is the set $S \cap T$ whose members are members of S *and* T. If two sets S and T have the property that $S \cap T = \emptyset$, then S and T are said to be ***disjoint***. For sets S and T their ***symmetric complement*** is the set

$$S \triangle T = (S - T) \cup (T - S).$$

Thus $S \triangle T$ is the set of objects in union $S \cup T$ that do not lie in the intersection $S \cap T$. The symmetric complement is so named because $S \triangle T = T \triangle S$. In Figure 1.1 we give Venn diagrams describing union, intersection, and symmetric complement.

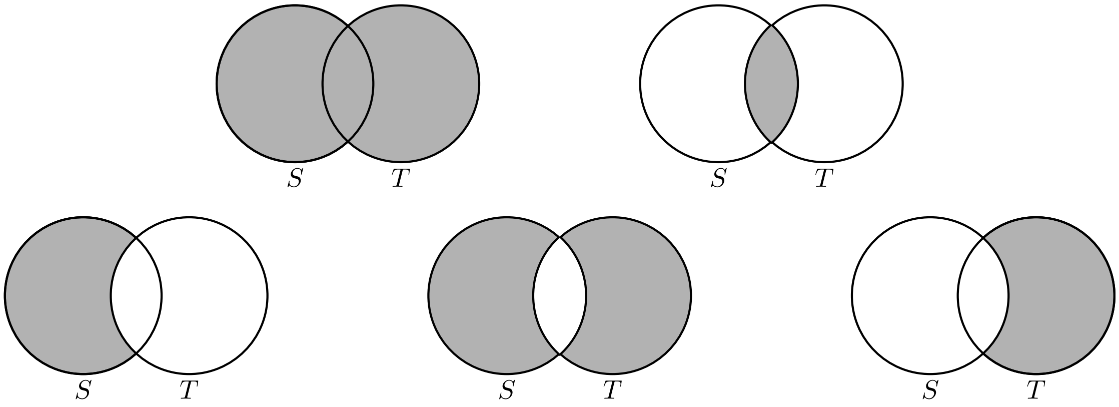


Figure 1.1 $S \cup T$ (top left), $S \cap T$ (top right), $S - T$ (bottom left), $S \triangle T$ (bottom middle), and $T - S$ (bottom right)

The following result gives some simple properties of pairwise unions and intersections of sets. We leave the straightforward verification of some or all of these to the reader as Exercise 1.1.5.

1.1.4 Proposition (Properties of unions and intersections) *For sets* S *and* T, *the following statements hold:*

(i) $S \cup \emptyset = S$;

(ii) $S \cap \emptyset = \emptyset$;

(iii) $S \cup S = S$;

(iv) $\mathrm{S} \cap \mathrm{S} = \mathrm{S}$;

(v) $\mathrm{S} \cup \mathrm{T} = \mathrm{T} \cup \mathrm{S}$ (***commutativity***);

(vi) $\mathrm{S} \cap \mathrm{T} = \mathrm{T} \cap \mathrm{S}$ (***commutativity***);

(vii) $\mathrm{S} \subseteq \mathrm{S} \cup \mathrm{T}$;

(viii) $\mathrm{S} \cap \mathrm{T} \subseteq \mathrm{S}$;

(ix) $\mathrm{S} \cup (\mathrm{T} \cup \mathrm{U}) = (\mathrm{S} \cup \mathrm{T}) \cup \mathrm{U}$ (***associativity***);

(x) $\mathrm{S} \cap (\mathrm{T} \cap \mathrm{U}) = (\mathrm{S} \cap \mathrm{T}) \cap \mathrm{U}$ (***associativity***);

(xi) $\mathrm{S} \cap (\mathrm{T} \cup \mathrm{U}) = (\mathrm{S} \cap \mathrm{T}) \cup (\mathrm{S} \cap \mathrm{U})$ (***distributivity***);

(xii) $\mathrm{S} \cup (\mathrm{T} \cap \mathrm{U}) = (\mathrm{S} \cup \mathrm{T}) \cap (\mathrm{S} \cup \mathrm{U})$ (***distributivity***).

We may more generally consider not just two sets, but an arbitrary collection $\mathscr{S}$ of sets. In this case we *posit* the existence of a set, called the ***union*** of the sets $\mathscr{S}$, with the property that it contains each element of each set $S \in \mathscr{S}$. Moreover, one can specify the subset of this big set to *only* contain members of sets from $\mathscr{S}$. This set we will denote by $\cup_{S\in\mathscr{S}} S$. We can also perform a similar construction with intersections of an arbitrary collection $\mathscr{S}$ of sets. Thus we denote by $\cap_{S\in\mathscr{S}} S$ the set, called the ***intersection*** of the sets $\mathscr{S}$, having the property that $x \in \cap_{S\in\mathscr{S}} S$ if $x \in S$ for every $S \in \mathscr{S}$. Note that we do not need to posit the existence of the intersection.

Let us give some properties of general unions and intersections as they relate to complements.

1.1.5 Proposition (De Morgan's[1] Laws) *Let* T *be a set and let* $\mathscr{S}$ *be a collection of subsets of* T. *Then the following statements hold:*

(i) $\mathrm{T} \setminus (\cup_{\mathrm{S}\in\mathscr{S}} \mathrm{S}) = \cap_{\mathrm{S}\in\mathscr{S}} (\mathrm{T} \setminus \mathrm{S})$;

(ii) $\mathrm{T} \setminus (\cap_{\mathrm{S}\in\mathscr{S}} \mathrm{S}) = \cup_{\mathrm{S}\in\mathscr{S}} (\mathrm{T} \setminus \mathrm{S})$.

Proof (i) Let $x \in T \setminus (\cup_{S\in\mathscr{S}})$. Then, for each $S \in \mathscr{S}$, $x \notin S$, or $x \in T \setminus S$. Thus $x \in \cap_{S\in\mathscr{S}}(T \setminus S)$. Therefore, $T \setminus (\cup_{S\in\mathscr{S}}) \supseteq \cap_{S\in\mathscr{S}}(T \setminus S)$. Conversely, if $x \in \cap_{S\in\mathscr{S}}(T \setminus S)$, then, for each $S \in \mathscr{S}$, $x \notin S$. Therefore, $x \notin \cup_{S\in\mathscr{S}}$. Therefore, $x \in T \setminus (\cup_{S\in\mathscr{S}})$, thus showing that $\cap_{S\in\mathscr{S}}(T \setminus S) \subseteq T \setminus (\cup_{S\in\mathscr{S}})$. It follows that $T \setminus (\cup_{S\in\mathscr{S}}) = \cap_{S\in\mathscr{S}}(T \setminus S)$.

(ii) This follows in much the same manner as part (i), and we leave the details to the reader. ■

1.1.6 Remark (Showing two sets are equal) Note that in proving part (i) of the preceding result, we proved two things. First we showed that $T \setminus (\cup_{S\in\mathscr{S}}) \subseteq \cap_{S\in\mathscr{S}}(T \setminus S)$ and then we showed that $\cap_{S\in\mathscr{S}}(T \setminus S) \subseteq T \setminus (\cup_{S\in\mathscr{S}})$. This is the standard means of showing that two sets are equal; first show that one is a subset of the other, and then show that the other is a subset of the one. •

For general unions and intersections, we also have the following generalisation of the distributive laws for unions and intersections. We leave the straightforward proof to the reader (Exercise 1.1.6)

[1] Augustus De Morgan (1806–1871) was a British mathematician whose principal mathematical contributions were to analysis and algebra.

1.1.7 Proposition (Distributivity laws for general unions and intersections) *Let* T *be a set and let* $\mathscr{S}$ *be a collection of sets. Then the following statements hold:*

(i) $\mathrm{T} \cap (\cup_{\mathrm{S}\in\mathscr{S}}\mathrm{S}) = \cup_{\mathrm{S}\in\mathrm{S}}(\mathrm{T} \cap \mathrm{S})$;

(ii) $\mathrm{T} \cup (\cap_{\mathrm{S}\in\mathscr{S}}\mathrm{S}) = \cap_{\mathrm{S}\in\mathrm{S}}(\mathrm{T} \cup \mathrm{S})$.

There is an alternative notion of the union of sets, one that retains the notion of membership in the original set. The issue that arises is this. If $S = \{1, 2\}$ and $T = \{2, 3\}$, then $S \cup T = \{1, 2, 3\}$. Note that we lose with the usual union the fact that 1 is an element of S only, but that 2 is an element of both S and T. Sometimes it is useful to retain these sorts of distinctions, and for this we have the following definition.

1.1.8 Definition (Disjoint union) ***missing stuff*** For sets S and T, their ***disjoint union*** is the set

$$S \overset{\circ}{\cup} T = \{(S, x) \mid x \in S\} \cup \{(T, y) \mid y \in T\}. \qquad \bullet$$

Let us see how the disjoint union differs from the usual union.

1.1.9 Example (Disjoint union) Let us again take the simple example $S = \{1, 2\}$ and $T = \{2, 3\}$. Then $S \cup T = \{1, 2, 3\}$ and

$$S \overset{\circ}{\cup} T = \{(S, 1), (S, 2), (T, 2), (T, 3)\}.$$

We see that the idea behind writing an element in the disjoint union as an ordered pair is that the first entry in the ordered pair simply keeps track of the set from which the element in the disjoint union was taken. In this way, if $S \cap T \neq \emptyset$, we are guaranteed that there will be no “collapsing” when the disjoint union is formed. •

1.1.3 Finite Cartesian products

As we have seen, if S is a set and if $x_1, x_2 \in S$, then $\{x_1, x_2\} = \{x_2, x_1\}$. There are times, however, when we wish to keep track of the order of elements in a set. To accomplish this and other objectives, we introduce the notion of an ordered pair. First, however, in order to make sure that we understand the distinction between ordered and unordered pairs, we make the following definition.

1.1.10 Definition (Unordered pair) If S is a set, an ***unordered pair*** from S is any subset of S with two elements. The collection of unordered pairs from S is denoted by $S^{(2)}$. •

Obviously one can talk about unordered collections of more than two elements of a set, and the collection of subsets of a set S comprised of k elements is denoted by $S^{(k)}$ and called the set of ***unordered* $\mathbf{k}$*-tuples***.

With the simple idea of an unordered pair, the notion of an ordered pair is more distinct.

1.1.11 Definition (Ordered pair and Cartesian product) Let S and T be sets, and let $x \in S$ and $y \in T$. The ***ordered pair*** of x and y is the set $(x, y) = \{\{x\}, \{x, y\}\}$. The ***Cartesian product*** of S and T is the set

$$S \times T = \{(x, y) \mid x \in S, \; y \in T\}. \qquad \bullet$$

The definition of the ordered pair seems odd at first. However, it is as it is to secure the objective that if two ordered pairs (x_1, y_1) and (x_2, y_2) are equal, then $x_1 = x_2$ and $y_1 = y_2$. The reader can check in Exercise 1.1.8 that this objective is in fact achieved by the definition. It is also worth noting that the form of the ordered pair as given in the definition is seldom used after its initial introduction.

Clearly one can define the Cartesian product of any finite number of sets $S_1, \dots, S_k$ inductively. Thus, for example, $S_1 \times S_2 \times S_3 = (S_1 \times S_2) \times S_3$. Note that, according to the notation in the definition, an element of $S_1 \times S_2 \times S_3$ should be written as $((x_1, x_2), x_3)$. However, it is immaterial that we define $S_1 \times S_2 \times S_3$ as we did, or as $S_1 \times S_2 \times S_3 = S_1 \times (S_2 \times S_3)$. Thus we simply write elements in $S_1 \times S_2 \times S_3$ as (x_1, x_2, x_3), and similarly for a Cartesian product $S_1 \times \cdots \times S_k$. The Cartesian product of a set with itself k-times is denoted by S^k. That is,

$$S^k = \underbrace{S \times \cdots \times S}_{k\text{-times}}.$$

In Section 1.4.2 we shall indicate how to define Cartesian products of more than finite collections of sets.

Let us give some simple examples.

1.1.12 Examples (Cartesian products)

1. If S is a set then note that $S \times \emptyset = \emptyset$. This is because there are no ordered pairs from S and $\emptyset$. It is just as clear that $\emptyset \times S = \emptyset$. It is also clear that, if $S \times T = \emptyset$, then either $S = \emptyset$ or $T = \emptyset$.
2. If $S = \{1, 2\}$ and $T = \{2, 3\}$, then

$$S \times T = \{(1, 2), (1, 3), (2, 2), (2, 3)\}. \qquad \bullet$$

Cartesian products have the following properties.

1.1.13 Proposition (Properties of Cartesian product) *For sets* S, T, U, *and* V, *the following statements hold:*

(i) $(\mathrm{S} \cup \mathrm{T}) \times \mathrm{U} = (\mathrm{S} \times \mathrm{U}) \cup (\mathrm{T} \times \mathrm{U})$;

(ii) $(\mathrm{S} \cap \mathrm{U}) \times (\mathrm{T} \cap \mathrm{V}) = (\mathrm{S} \times \mathrm{T}) \cap (\mathrm{U} \times \mathrm{V})$;

(iii) $(\mathrm{S} - \mathrm{T}) \times \mathrm{U} = (\mathrm{S} \times \mathrm{U}) - (\mathrm{T} \times \mathrm{U})$.

Proof Let us prove only the first identity, leaving the remaining two to the reader. Let $(x, u) \in (S \cup T) \times U$. Then $x \in S \cup T$ and $u \in U$. Therefore, x is an element of at least one of S and T. Without loss of generality, suppose that $x \in S$. Then $(x, u) \in S \times U$ and so $(x, u) \in (S \times U) \cup (T \times U)$. Therefore, $(S \cup T) \times U = (S \times U) \cup (T \times U)$. Conversely, suppose that $(x, u) \in (S \times U) \cup (T \times U)$. Without loss of generality, suppose that $(x, u) \in S \times U$. Then $x \in S \subseteq S \cup T$ and $u \in U$. Therefore, $(x, u) \in (S \cup T) \times U$. Thus $(S \times U) \cup (T \times U) \subseteq (S \cup T) \times U$, giving the result. ■

1.1.14 Remark (“Without loss of generality”) In the preceding proof, we twice employed the expression “without loss of generality.” This is a commonly encountered expression, and is frequently used in one of the following two contexts. The first, as above, indicates that one is making an arbitrary selection, but that were another arbitrary selection to have been made, the same argument holds. This is a more or less straightforward use of “without loss of generality.” A more sophisticated use of the expression might indicate that one is making a simplifying assumption, and that this is okay, because it can be shown that the general case follows easily from the simpler one. The trick is to then understand *how* the general case follows from the simpler one, and this can sometimes be nontrivial, depending on the willingness of the writer to describe this process. •

Exercises

1.1.1 Prove that the empty set is a subset of every set.
***Hint:** Assume the converse and arrive at an absurdity.*

1.1.2 Let S be a set, let $A, B, C \subseteq S$, and let $\mathscr{A}, \mathscr{B} \subseteq \mathbf{2}^S$.

(a) Show that $A \triangle \emptyset = A$.

(b) Show that $(S \setminus A) \triangle (S \setminus B) = A \triangle B$.

(c) Show that $A \triangle C \subseteq (A \triangle B) \cup (B \triangle C)$.

(d) Show that

$$\Bigl(\cup_{A\in\mathscr{A}} A\Bigr) \triangle \Bigl(\cup_{B\in\mathscr{B}} B\Bigr) \subseteq \cup_{(A,B)\in\mathscr{A}\times\mathscr{B}} (A \triangle B),$$
$$\Bigl(\cap_{A\in\mathscr{A}} A\Bigr) \triangle \Bigl(\cap_{B\in\mathscr{B}} B\Bigr) \subseteq \cap_{(A,B)\in\mathscr{A}\times\mathscr{B}} (A \triangle B),$$
$$\cap_{(A,B)\in\mathscr{A}\times\mathscr{B}} (A \triangle B) \subseteq \Bigl(\cap_{A\in\mathscr{A}} A\Bigr) \triangle \Bigl(\cup_{B\in\mathscr{B}} B\Bigr).$$

1.1.3 If S is a set with n members, show that $\mathbf{2}^S$ is a set with 2^n members.

1.1.4 Let S be a set with m elements. Show that the number of subsets of S having k distinct elements is $\binom{m}{k} = \frac{m!}{k!(m-k)!}$.

1.1.5 Prove as many parts of Proposition 1.1.4 as you wish.

1.1.6 Prove Proposition 1.1.7.

1.1.7 Let S be a set with n members and let T be a set with m members. Show that $S \overset{\circ}{\cup} T$ is a set with nm members.

1.1.8 Let S and T be sets, let $x_1, x_2 \in S$, and let $y_1, y_2 \in T$. Show that $(x_1, y_1) = (x_2, y_2)$ if and only if $x_1 = x_2$ and $y_1 = y_2$.

Section 1.2

Relations

Relations are a fundamental ingredient in the description of many mathematical ideas. One of the most valuable features of relations is that they allow many useful constructions to be explicitly made only using elementary ideas from set theory.

Do I need to read this section? The ideas in this section will appear in many places in the series, so this material should be regarded as basic. However, readers looking to proceed with minimal background can skip the section, referring back to it when needed. •

1.2.1 Definitions

We shall describe in this section "binary relations," or relations between elements of two sets. It is possible to define more general sorts of relations where more sets are involved. However, these will not come up for us.

1.2.1 **Definition (Relation)** A ***binary relation from* S *to* T** (or simply a ***relation from* S *to* T**) is a subset of $S \times T$. If $R \subseteq S \times T$ and if $(x, y) \in R$, then we shall write $x\,R\,y$, meaning that x and y are related by R. A relation from S to S is a ***relation in* S**. •

The definition is simple. Let us give some examples to give it a little texture.

1.2.2 **Examples (Relations)**

1. Let S be the set of husbands and let T be the set of wives. Define a relation R from S to T by asking that $(x, y) \in R$ if x is married to y. Thus, to say that x and y are related in this case means to say that x is married to y.
2. Let S be a set and consider the relation R in the power set $\mathbf{2}^S$ of S given by
$$R = \{(A, B) \mid A \subseteq B\}.$$
Thus A is related to B if A is a subset of B.
3. Let S be a set and define a relation R in S by
$$R = \{(x, x) \mid x \in S\}.$$
Thus, under this relation, two members in S are related if and only if they are equal.
4. Let S be the set of integers, let k be a positive integer, and define a relation R_k in S by
$$R_k = \{(n_1, n_2) \mid n_1 - n_2 = k\}.$$
Thus, if $n \in S$, then all integers of the form $n + mk$ for an integer m are related to n. •

1.2.3 Remark ("If" versus "if and only if") In part 3 of the preceding example we used the expression "if and only if" for the first time. It is, therefore, worth saying a few words about this commonly used terminology. One says that statement A holds "if and only if" statement B holds to mean that statements A and B are exactly equivalent. Typically, this language arises in theorem statements. In proving such theorems, it is important to note that one must prove *both* that statement A implies statement B *and* that statement B implies statement A.

To confuse matters, when stating a definition, the convention is to use "if" rather than "if and only if". It is not uncommon to see "if and only if" used in definitions, the thinking being that a definition makes the thing being defined as equivalent to what it is defined to be. However, there is a logical flaw here. Indeed, suppose one is defining "X" to mean that "Proposition A applies". If one writes "X if and only if Proposition A applies" then this makes no sense. Indeed the "only if" part of this statement says that the statement "Proposition A applies" if "X" holds. But "X" is undefined except by saying that it holds when "Proposition A applies". •

In the next section we will encounter the notion of the inverse of a function; this idea is perhaps known to the reader. However, the notion of inverse also applies to the more general setting of relations.

1.2.4 Definition (Inverse of a relation) If $R \subseteq S \times T$ is a relation from S to T, then the ***inverse*** of R is the relation R^{-1} from T to S defined by

$$R^{-1} = \{(y, x) \in T \times S \mid (x, y) \in R\}. \qquad \bullet$$

There are a variety of properties that can be bestowed upon relations to ensure they have certain useful attributes. The following is a partial list of such properties.

1.2.5 Definition (Properties of relations) Let S be a set and let R be a relation in S. The relation R is:

(i) ***reflexive*** if $(x, x) \in R$ for each $x \in S$;

(ii) ***irreflexive*** if $(x, x) \notin R$ for each $x \in S$;

(iii) ***symmetric*** if $(x_1, x_2) \in R$ implies that $(x_2, x_1) \in R$;

(iv) ***antisymmetric*** if $(x_1, x_2) \in R$ and $(x_2, x_1) \in R$ implies that $x_1 = x_2$;

(v) ***transitive*** if $(x_1, x_2) \in R$ and $(x_2, x_3) \in R$ implies that $(x_1, x_3) \in R$. •

1.2.6 Examples (Example 1.2.2 cont'd)

1. The relation of inclusion in the power set $\mathbf{2}^S$ of a set S is reflexive, antisymmetric, and transitive.
2. The relation of equality in a set S is reflexive, symmetric, antisymmetric, and transitive.
3. The relation R_k in the set S of integers is reflexive, symmetric, and transitive. •

1.2.2 Equivalence relations

In this section we turn our attention to an important class of relations, and we indicate why these are important by giving them a characterisation in terms of a decomposition of a set.

1.2.7 Definition (Equivalence relation, equivalence class) An ***equivalence relation*** in a set S is a relation R that is reflexive, symmetric, and transitive. For $x \in S$, the set of elements of S related to x is denoted by $[x]$, and is the ***equivalence class*** of x with respect to R. An element x' in an equivalence class $[x]$ is a ***representative*** of that equivalence class. The set of equivalence classes is denoted by S/R (typically pronounced as **S *modulo* R**). •

It is common to denote that two elements $x_1, x_2 \in S$ are related by an equivalence relation by writing $x_1 \sim x_2$. Of the relations defined in Example 1.2.2, we see that those in parts 3 and 4 are equivalence relations, but that in part 2 is not.

Let us now characterise equivalence relations in a more descriptive manner. We begin by defining a (perhaps seemingly unrelated) notion concerning subsets of a set.

1.2.8 Definition (Partition of a set) A ***partition*** of a set S is a collection $\mathscr{A}$ of subsets of S having the properties that

(i) two distinct subsets in $\mathscr{A}$ are disjoint and

(ii) $S = \cup_{A \in \mathscr{A}} A$. •

We now prove that there is an exact correspondence between equivalence classes associated to an equivalence relation.

1.2.9 Proposition (Equivalence relations and partitions) *Let* S *be a set and let* R *be an equivalence relation in* S*. Then the set of equivalence classes with respect to* R *is a partition of* S*.*

Conversely, if $\mathscr{A}$ is a partition of S*, then the relation*

$$\{(x_1, x_2) \mid x_1, x_2 \in A \textit{ for some } A \in \mathscr{A}\}$$

is an equivalence relation in S*.*

Proof We first claim that two distinct equivalence classes are disjoint. Thus we let $x_1, x_2 \in S$ and suppose that $[x_1] \neq [x_2]$. Suppose that $x \in [x_1] \cap [x_2]$. Then $x \sim x_1$ and $x \sim x_2$, or, by transitivity of R, $x_1 \sim x$ and $x \sim x_2$. By transitivity of R, $x_1 \sim x_2$, contradicting the fact that $[x_1] \neq [x_2]$. To show that S is the union of its equivalence classes, merely note that, for each $x \in S$, $x \in [x]$ by reflexivity of R.

Now let $\mathscr{A}$ be a partition and defined R as in the statement of the proposition. Let $x \in S$ and let A be the element of $\mathscr{A}$ that contains x. Then clearly we see that $(x, x) \in R$ since $x \in A$. Thus R is reflexive. Next let $(x_1, x_2) \in R$ and let A be the element of $\mathscr{A}$ such that $x_1, x_2 \in A$. Clearly then, $(x_2, x_1) \in R$, so R is symmetric. Finally, let $(x_1, x_2), (x_2, x_3 \in R$. Then there are elements $A_{12}, A_{23} \in \mathscr{A}$ such that $x_1, x_2 \in A_{12}$ and such that $x_2, x_3 \in A_{23}$. Since A_{12} and A_{23} have the point x_2 in common, we must have $A_{12} = A_{23}$. Thus $(x_1, x_3 \in A_{12} = A_{23}$, giving transitivity of R. ■

Exercises

1.2.1 In a set S define a relation $R = \{(x, y) \in S \times S \mid x = y\}$.

(a) Show that R is an equivalence relation.

(b) Show that $S/R = S$.

Section 1.3

Maps

Another basic concept in all of mathematics is that of a map between sets. Indeed, many of the interesting objects in mathematics are maps of some sort. In this section we review the notation associated with maps, and give some simple properties of maps.

Do I need to read this section? The material in this section is basic, and will be used constantly throughout the series. Unless you are familiar already with maps and the notation associated to them, this section is essential reading. •

1.3.1 Definitions and notation

We begin with the definition.

1.3.1 **Definition (Map)** For sets S and T, a ***map*** from S to T is a relation R from S to T having the property that, for each $x \in S$, there exists a unique $y \in T$ such that $(x, y) \in R$. The set S is the ***domain*** of the map and the set T is the ***codomain*** of the map. The set of maps from S to T is denoted by T^S.[2] •

By definition, a map is a relation. This is not how one most commonly thinks about a map, although the definition serves to render the concept of a map in terms of concepts we already know. Suppose one has a map from S to T defined by a relation R. Then, given $x \in S$, there is a single $y \in T$ such that x and y are related. Denote this element of T by $f(x)$, since it is defined by x. When one refers to a map, one more typically refers to the assignment of the element $f(x) \in T$ to $x \in S$. Thus one refers to the map as f, leaving aside the baggage of the relation as in the definition. Indeed, this is how we from now on will think of maps. The definition above does, however, have some use, although we alter our language, since we are now thinking of a map as an "assignment." We call the set

$$\operatorname{graph}(f) = \{(x, f(x)) \mid x \in S\} \subseteq S \times T$$

(which we originally called the map in Definition 1.3.1) the ***graph*** of the map $f\colon S \to T$.

If one wishes to indicate a map f with domain S and codomain T, one typically writes $f\colon S \to T$ to compactly express this. If one wishes to *define* a map by saying what it does, the notation

$$\begin{aligned} f\colon S &\to T \\ x &\mapsto \text{what } x \text{ gets mapped to} \end{aligned}$$

[2] The idea behind this notation is the following. A map from S to T assigns to each point in S a point in T. If S and T are finite sets with k and l elements, respectively, then there are l possible values that can be assigned to each of the k elements of S. Thus the set of maps has l^k elements.

is sometimes helpful. Sometimes we shall write this in the text as $f\colon x \mapsto$ "what x gets mapped to". Note the distinct uses of the symbols "$\to$" and "$\mapsto$".

1.3.2 Notation (f versus f(x)) Note that a map is denoted by "f". It is quite common to see the expression "consider the map $f(x)$". Taken literally, these words are difficult to comprehend. First of all, x is unspecified. Second of all, even if x were specified, $f(x)$ is an element of T, not a map. Thus it is considered bad form mathematically to use an expression like "consider the map $f(x)$". However, there are times when it is quite convenient to use this poor notation, with an understanding that some compromises are being made. For instance, in this volume, we will be frequently dealing simultaneously with functions of both time (typically denoted by t) and frequency (typically denoted by ν). Thus it would be convenient to write "consider the map $f(t)$" when we wish to write a map that we are considering as a function of time, and similarly for frequency. Nonetheless, we shall refrain from doing this, and shall consistently use the mathematically precise language "consider the map f". •

The following is a collection of examples of maps. Some of these examples are not just illustrative, but also define concepts and notation that we will use throughout the series.

1.3.3 Examples (Maps)

1. There are no maps having $\emptyset$ as a domain or codomain since there are no elements in the empty set.
2. If S is a set and if $T \subseteq S$, then the map $i_T\colon T \to S$ defined by $i_T(x) = x$ is called the ***inclusion*** of T in S.
3. The inclusion map $i_S\colon S \to S$ of a set S into itself (since $S \subseteq S$) is the ***identity map***, and we denote it by id_S.
4. If $f\colon S \to T$ is a map and if $A \subseteq S$, then the map from A to T which assigns to $x \in A$ the value $f(x) \in T$ is called the ***restriction*** of f to A, and is denoted by $f|A\colon A \to T$.
5. If S is a set with $A \subseteq S$, then the map χ_A from S to the integers defined by
$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$
is the ***characteristic function*** of A.
6. If $S_1, \dots, S_k$ are sets, if $S_1 \times \cdots \times S_k$ is the Cartesian product, and if $j \in \{1, \dots, k\}$, then the map
$$\begin{aligned} \mathrm{pr}_j\colon S_1 \times \cdots \times S_j \times \cdots \times S_k &\to S_j \\ (x_1, \dots, x_j, \dots, x_k) &\mapsto x_j \end{aligned}$$
is the ***projection onto the* j*th factor***.
7. If R is an equivalence relation in a set S, then the map $\pi_R\colon S \to S/R$ defined by $\pi_R(x) = [x]$ is called the ***canonical projection*** associated to R.

8. If S, T, and U are sets and if $f\colon S \to T$ and $g\colon T \to U$ are maps, then we define a map $g \circ f\colon S \to U$ by $g \circ f(x) = g(f(x))$. This is the ***composition*** of f and g.
9. If S and $T_1, \dots, T_k$ are sets then a map $f\colon S \to T_1 \times \cdots \times T_k$ can be written as

$$f(x) = (f_1(x), \dots, f_k(x))$$

for maps $f_j\colon S \to T_j$, $j \in \{1, \dots, k\}$. In this case we will write $f = f_1 \times \cdots \times f_k$. •

Next we introduce the notions of images and preimages of points and sets.

1.3.4 Definition (Image and preimage) Let S and T be sets and let $f\colon S \to T$ be a map.

(i) If $A \subseteq S$, then $f(A) = \{f(x) \mid x \in A\}$.

(ii) The ***image*** of f is the set $\mathrm{image}(f) = f(S) \subseteq T$.

(iii) If $B \subseteq T$, then $f^{-1}(B) = \{x \in S \mid f(x) \in B\}$ is the ***preimage*** of B under f. If $B = \{y\}$ for some $y \in T$, then we shall often write $f^{-1}(y)$ rather that $f^{-1}(\{y\})$. •

Note that one can think of f as being a map from $\mathbf{2}^S$ to $\mathbf{2}^T$ and of f^{-1} as being a map from $\mathbf{2}^T$ to $\mathbf{2}^S$. Here are some elementary properties of f and f^{-1} thought of in this way.

1.3.5 Proposition (Properties of images and preimages) *Let* S *and* T *be sets, let* $\mathrm{f}\colon \mathrm{S} \to \mathrm{T}$ *be a map, let* $\mathrm{A} \subseteq \mathrm{S}$ *and* $\mathrm{B} \subseteq \mathrm{T}$*, and let* $\mathscr{A}$ *and* $\mathscr{B}$ *be collections of subsets of* S *and* T*, respectively. Then the following statements hold:*

(i) $\mathrm{A} \subseteq \mathrm{f}^{-1}(\mathrm{f}(\mathrm{A}))$;

(ii) $\mathrm{f}(\mathrm{f}^{-1}(\mathrm{B})) \subseteq \mathrm{B}$;

(iii) $\cup_{\mathrm{A}\in\mathscr{A}}\mathrm{f}(\mathrm{A}) = \mathrm{f}(\cup_{\mathrm{A}\in\mathscr{A}}\mathrm{A})$;

(iv) $\cup_{\mathrm{B}\in\mathscr{B}}\mathrm{f}^{-1}(\mathrm{B}) = \mathrm{f}^{-1}(\cup_{\mathrm{B}\in\mathscr{B}}\mathrm{B})$;

(v) $\cap_{\mathrm{A}\in\mathscr{A}}\mathrm{f}(\mathrm{A}) = \mathrm{f}(\cap_{\mathrm{A}\in\mathscr{A}}\mathrm{A})$;

(vi) $\cap_{\mathrm{B}\in\mathscr{B}}\mathrm{f}^{-1}(\mathrm{B}) = \mathrm{f}^{-1}(\cap_{\mathrm{B}\in\mathscr{B}}\mathrm{B})$.

Proof We shall prove only some of these, leaving the remainder for the reader to complete.

(i) Let $x \in A$. Then $x \in f^{-1}(f(x))$ since $f(x) = f(x)$.

(iii) Let $y \in \cup_{A\in\mathscr{A}} f(A)$. Then $y = f(x)$ for some $x \in \cup_{A\in\mathscr{A}} A$. Thus $y \in f(\cup_{A\in\mathscr{A}} A)$. Conversely, let $y \in f(\cup_{A\in\mathscr{A}} A)$. Then, again, $y = f(x)$ for some $x \in \cup_{A\in\mathscr{A}} A$, and so $y \in \cup_{A\in\mathscr{A}} f(A)$.

(vi) Let $x \in \cap_{B\in\mathscr{B}} f^{-1}(B)$. Then, for each $B \in \mathscr{B}$, $x \in f^{-1}(B)$. Thus $f(x) \in B$ for all $B \in \mathscr{B}$ and so $f(x) \in \cap_{B\in\mathscr{B}} B$. Thus $x \in f^{-1}(\cap_{B\in\mathscr{B}} B)$. Conversely, if $x \in f^{-1}(\cap_{B\in\mathscr{B}} B)$, then $f(x) \in B$ for each $B \in \mathscr{B}$. Thus $x \in f^{-1}(B)$ for each $B \in \mathscr{B}$, or $x \in \cap_{B\in\mathscr{B}} f^{-1}(B)$. ■

1.3.2 Properties of maps

Certain basic features of maps will be of great interest.

1.3.6 Definition (Injection, surjection, bijection) Let S and T be sets. A map $f\colon S \to T$ is:

(i) ***injective***, or an ***injection***, if $f(x) = f(y)$ implies that $x = y$;

(ii) ***surjective***, or a ***surjection***, if $f(S) = T$;

(iii) ***bijective***, or a ***bijection***, if it is both injective and surjective. •

1.3.7 Remarks (One-to-one, onto, 1–1 correspondence)

1. It is not uncommon for an injective map to be said to be ***1–1*** or ***one-to-one***, and that a surjective map be said to be ***onto***. In this series, we shall exclusively use the terms injective and surjective, however. These words appear to have been given prominence by their adoption by Bourbaki (see footnote on page iv).
2. If there exists a bijection $f\colon S \to T$ between sets S and T, it is common to say that there is a ***1–1 correspondence*** between S and T. This can be confusing if one is familiar with the expression "1–1" as referring to an injective map. The words "1–1 correspondence" mean that there is a bijection, not an injection. In case S and T are in 1–1 correspondence, we shall also say that S and T are ***equivalent***. •

Closely related to the above concepts, although not immediately obviously so, are the following notions of inverse.

1.3.8 Definition (Left-inverse, right-inverse, inverse) Let S and T be sets, and let $f\colon S \to T$ be a map. A map $g\colon T \to S$ is:

(i) a ***left-inverse*** of f if $g \circ f = \mathrm{id}_S$;

(ii) a ***right-inverse*** of f if $f \circ g = \mathrm{id}_T$;

(iii) an ***inverse*** of f if it is both a left- and a right-inverse. •

In Definition 1.2.4 we gave the notion of the inverse of a relation. Functions, being relations, also possess inverses in the sense of relations. We ask the reader to explore the relationships between the two concepts of inverse in Exercise 1.3.7.

The following result relates these various notions of inverse to the properties of injective, surjective, and bijective.

1.3.9 Proposition (Characterisation of various inverses) *Let* S *and* T *be sets and let* $\mathrm{f}\colon \mathrm{S} \to \mathrm{T}$ *be a map. Then the following statements hold:*

(i) f *is injective if and only if it possesses a left-inverse;*

(ii) f *is surjective if and only if it possess a right-inverse;*

(iii) f *is bijective if and only if it possesses an inverse;*

(iv) there is at most one inverse for f*;*

(v) if f *possesses a left-inverse and a right-inverse, then these necessarily agree.*

Proof (i) Suppose that f is injective. For $y \in \mathrm{image}(f)$, define $g(y) = x$ where $f^{-1}(y) = \{x\}$, this being well-defined since f is injective. For $y \notin \mathrm{image}(f)$, define $g(y) = x_0$ for some $x_0 \in S$. The map g so defined is readily verified to satisfy $g \circ f = \mathrm{id}_S$, and so is a left-inverse. Conversely, suppose that f possesses a left-inverse g, and let $x_1, x_2 \in S$ satisfy $f(x_1) = f(x_2)$. Then $g \circ f(x_1) = g \circ f(x_2)$, or $x_1 = x_2$. Thus f is injective.

(ii) Suppose that f is surjective. For $y \in T$ let $x \in f^{-1}(y)$ and define $g(y) = x$.[3] With g so defined it is easy to see that $f \circ g = \mathrm{id}_T$, so that g is a right-inverse. Conversely, suppose that f possesses a right-inverse g. Now let $y \in T$ and take $x = g(y)$. Then $f(x) = f \circ g(y) = y$, so that f is surjective.

(iii) Since f is bijective, it possesses a left-inverse g_L and a right-inverse g_R. We claim that these are equal, and each is actually an inverse of f. We have

$$g_L = g_L \circ \mathrm{id}_T = g_L \circ f \circ g_R = \mathrm{id}_S \circ g_R = g_R,$$

showing equality of g_L and g_R. Thus each is a left- and a right-inverse, and therefore an inverse for f.

(iv) Let g_1 and g_2 be inverses for f. Then, just as in part (iii),

$$g_1 = g_1 \circ \mathrm{id}_T = g_1 \circ f \circ g_2 = \mathrm{id}_S \circ g_2 = g_2.$$

(v) This follows from the proof of part (iv), noting that there we only used the facts that g_1 is a left-inverse and that g_2 is a right-inverse. ■

In Figure 1.2 we depict maps that have various of the properties of injectivity, surjectivity, or bijectivity. From these cartoons, the reader may develop some intuition for Proposition 1.3.9. In the case that $f\colon S \to T$ is a bijection, we denote its unique inverse by $f^{-1}\colon T \to S$. The confluence of the notation f^{-1} introduced when discussing preimages is not a problem, in practice.

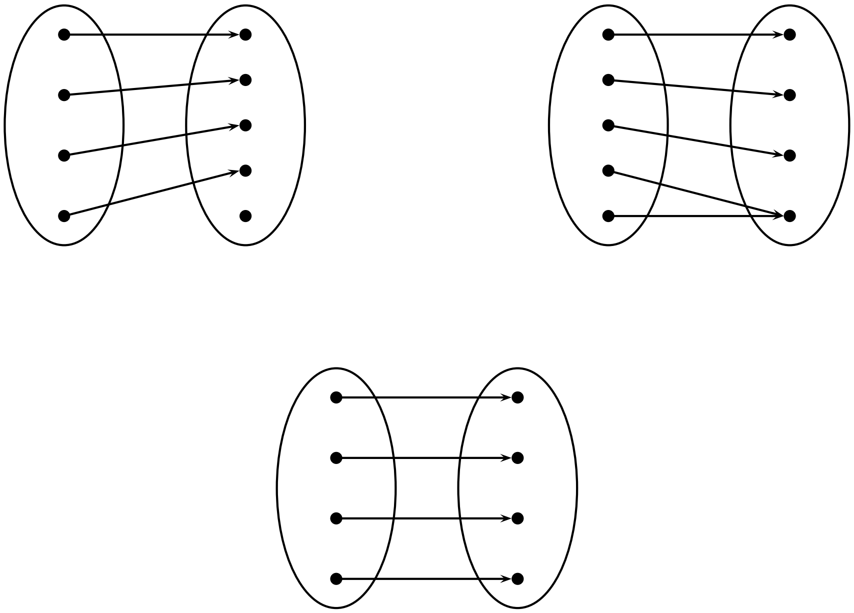


Figure 1.2 A depiction of maps that are injective but not surjective (top left), surjective but not injective (top right), and bijective (bottom)

[3] Note that the ability to choose an x from each set $f^{-1}(y)$ requires the Axiom of Choice (see Section **??**).

It is worth mentioning at this point that the characterisation of left- and right-inverses in Proposition 1.3.9 is not usually very helpful. Normally, in a given setting, one will want these inverses to have certain properties. For vector spaces, for example, one may want left- or right-inverses to be linear (see ***missing stuff***), and for topological spaces, for another example, one may want a left- or right-inverse to be continuous (see Chapter **??**).

1.3.3 Graphs and commutative diagrams

Often it is useful to be able to understand the relationship between a number of maps by representing them together in a diagram. We shall be somewhat precise about what we mean by a diagram by making it a special instance of a graph. We shall encounter graphs in ***missing stuff***, although for the present purposes we merely use them as a means of making precise the notion of a commutative diagram.

First the definitions for graphs.

1.3.10 Definition (Graph) A ***graph*** is a pair (V, E) where V is a set, an element of which is called a ***vertex***, and E is a subset of the set $V^{(2)}$ of unordered pairs from V, an element of which is called an ***edge***. If $\{v_1, v_2\} \in E$ is an edge, then the vertices v_1 and v_2 are the ***endvertices*** of this edge. •

In a graph, it is the way that vertices and edges are related that is of interest. To capture this structure, the following language is useful.

1.3.11 Definition (Adjacent and incident) Let (V, E) be a graph. Two vertices $v_1, v_2 \in V$ are ***adjacent*** if $\{v_1, v_2\} \in E$ and a vertex $v \in V$ and an edge $e \in E$ are ***incident*** if there exists $v' \in V$ such that $e = \{v, v'\}$. •

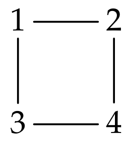
One typically represents a graph by placing the vertices in some sort of array on the page, and then drawing a line connecting two vertices if there is a corresponding edge associated with the two vertices. Some examples make this process clear.

1.3.12 Examples (Graphs)

1. Consider the graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{\{1, 2\}, \{1, 3\}, \{2, 4\}, \{3, 4\}\}.$$

There are many ways one can lay out the vertices on the page, but for this diagram, it is most convenient to arrange them in a square. Doing so gives rise to the following representation of the graph:

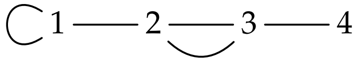


The vertices 1 and 2 are adjacent, but the vertices 1 and 4 are not. The vertex 1 and the edge $\{1, 2\}$ are incident, but the vertex 1 and the edge $\{3, 4\}$ are not.

2. For the graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{\{1, 2\}, \{2, 3\}, \{2, 3\}, \{3, 4\}\}$$

we have the representation



Note that we allow the same edge to appear twice, and we allow for an edge to connect a vertex to itself. We observe that the vertices 2 and 3 are adjacent, but the vertices 1 and 3 are not. Also, the vertex 3 and the edge $\{2, 3\}$ are incident, but the vertex 4 and the edge $\{1, 2\}$ are not. •

Often one wishes to attach "direction" to vertices. This is done with the following notion.

1.3.13 Definition (Directed graph) A ***directed graph***, or ***digraph***, is a pair (V, E) where V is a set an element of which is called a ***vertex*** and E is a subset of the set $V \times V$ of ordered pairs from V an element of which is called an ***edge***. If $e = (v_1, v_2) \in E$ is an edge, then v_1 is the ***source*** for e and v_2 is the ***target*** for e. •

Note that every directed graph is certainly also a graph, since one can assign an unordered pair to every ordered pair of vertices.

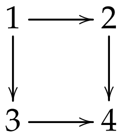
The examples above of graphs are easily turned into directed graphs, and we see that to represent a directed graph one needs only to put a "direction" on an edge, typically via an arrow.

1.3.14 Examples (Directed graphs)

1. Consider the directed graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{(1, 2), (1, 3), (2, 4), (3, 4)\}.$$

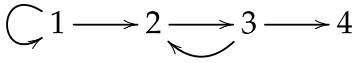
A convenient representation of this directed graph is as follows:



2. For the directed graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{(1, 1), (1, 2), (2, 3), (2, 3), (3, 4)\}$$

we have the representation



•

Of interest in graph theory is the notion of connecting two, perhaps nonadjacent, vertices with a sequence of edges (the notion of a sequence is familiar, but will be made precise in Section 1.4.3). This is made precise as follows.

1.3.15 Definition (Path)

(i) If (V, E) is a graph, a ***path*** in the graph is a sequence $(a_j)_{j\in\{1,\dots,k\}}$ in $V \cup E$ with the following properties:

  (a) $a_1, a_k \in V$;

  (b) for $j \in \{1, \dots, k-1\}$, if $a_j \in V$ (resp. $a_j \in E$), then $a_{j+1} \in E$ (resp. $a_{j+1} \in V$).

(ii) If (V, E) is a directed graph, a ***path*** in the graph is a sequence $(a_j)_{j\in\{1,\dots,k\}}$ in $V \cup E$ with the following properties:

  (a) $(a_j)_{j\in\{1,\dots,k\}}$ is a path in the graph associated to (V, E);

  (b) for $j \in \{2, \dots, k-1\}$, if $a_j \in E$, then $a_j = (a_{j-1}, a_{j+1})$.

(iii) If $(a_j)_{j\in\{1,\dots,k\}}$ is a path, the ***length*** of the path is the number of edges in the path.

(iv) For a path $(a_j)_{j\in\{1,\dots,k\}}$, the ***source*** is the vertex a_1 and the ***target*** is the vertex a_k. •

Let us give some examples of paths for graphs and for directed graphs.

1.3.16 Examples (Paths)

1. For the graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{\{1, 2\}, \{1, 3\}, \{2, 4\}, \{3, 4\}\},$$

there are an infinite number of paths. Let us list a few:

(a) (1), (2), (3), and (4);

(b) $(4, \{3, 4\}, 3, \{1, 3\}, 1)$;

(c) $(1, \{1, 2\}, 2, \{2, 4\}, 4, \{3, 4\}, 3, \{1, 3\}, 1)$;

(d) $(1, \{1, 2\}, 2, \{1, 2\}, 1, \{1, 2\}, 2, \{1, 2\}, 1)$.

Note that for this graph there are infinitely many paths.

2. For the directed graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{(1, 2), (1, 3), (2, 4), (3, 4)\},$$

there are a finite number of paths:

(a) (1), (2), (3), and (4);

(b) $(1, (1, 2), 2)$;

(c) $(1, (1, 2), 2, (2, 4), 4)$;

(d) $(1, (1, 3), 3)$;

(e) $(1, (1, 3), 3, (2, 4), 4)$;

(f) $(2, (2, 4))$;

(g) $(3, (3, 4), 4)$.

3. For the graph (V, E) with

$$V = \{1, 2, 3, 4\}, \quad E = \{\{1, 2\}, \{2, 3\}, \{2, 3\}, \{3, 4\}\}$$

some examples of paths are:

   (a) (1), (2), (3), and (4);
   (b) $(1, \{1,2\}, 2, \{2,3\}, 3, \{2,3\}, 2, \{1,2\}, 1)$;
   (c) $(4, \{3,4\}, 3)$.

   There are an infinite number of paths for this graph.

4. For the directed graph (V, E) with

$$V = \{1,2,3,4\}, \quad E = \{(1,1),(1,2),(2,3),(2,3),(3,4)\}$$

   some paths include:

   (a) (1), (2), (3), and (4);
   (b) $(1,(1,2),2,(2,3),3,(3,2),2,(2,3),3,(3,4),4)$;
   (c) $(3,(3,4),4)$.

   This directed graph has an infinite number of paths by virtue of the fact that the path $(2,(2,3),3,(3,2),2)$ can be repeated an infinite number of times. •

1.3.17 Notation (Notation for paths of nonzero length) For paths which contain at least one edge, i.e., which have length at least 1, the vertices in the path are actually redundant. For this reason we will often simply write a path as the sequence of edges contained in the path, since the vertices can be obviously deduced. •

There is a great deal one can say about graphs, a little of which we will say in ***missing stuff***. However, for our present purposes of defining diagrams, the notions at hand are sufficient. In the definition we employ Notation 1.3.17.

1.3.18 Definition (Diagram, commutative diagram) Let (V, E) be a directed graph.

(i) A ***diagram*** on (V, E) is a family $(S_v)_{v\in V}$ of sets associated with each vertex and a family $(f_e)_{e\in E}$ of maps associated with each edge such that, if $e = (v_1, v_2)$, then f_e has domain S_{v_1} and codomain S_{v_2}.

(ii) If $P = (e_j)_{j\in\{1,\dots,k\}}$ is a path of nonzero length in a diagram on (V, E), the ***composition*** along P is the map $f_{e_k} \circ \cdots \circ f_{e_1}$.

(iii) A diagram is ***commutative*** if, for every two vertices $v_1, v_2 \in V$ and any two paths P_1 and P_2 with source v_1 and target v_2, the composition along P_1 is equal to the composition along P_2. •

The notion of a diagram, and in particular a commutative diagram is straightforward.

1.3.19 Examples (Diagrams and commutative diagrams)

1. Let S_1, S_2, S_3, and S_4 be sets and consider maps $f_{21}\colon S_1 \to S_2$, $f_{31}\colon S_1 \to S_3$, $f_{42}\colon S_2 \to S_4$, and $f_{43}\colon S_3 \to S_4$.[4] ***missing stuff*** Note that if we assign set S_j to j for each $j \in \{1,2,3,4\}$, then this gives a diagram on (V, E) where

$$V = \{1,2,3,4\}, \quad E = \{(1,2),(1,3),(2,4),(3,4)\}.$$

[4] It might seem more natural to write, for example, $f_{12}\colon S_1 \to S_2$ to properly represent the normal order of the domain and codomain. However, we instead write $f_{21}\colon S_1 \to S_2$ for reasons having to do with conventions that will become convenient in .

This diagram can be represented by

$$\begin{tikzcd} S_1 \arrow[r, "f_{21}"] \arrow[d, "f_{31}"'] & S_2 \arrow[d, "f_{42}"] \\ S_3 \arrow[r, "f_{43}"'] & 4 \end{tikzcd}$$

The diagram is commutative if and only if $f_{42} \circ f_{21} = f_{43} \circ f_{31}$.

2. Let S_1, S_2, S_3, and S_4 be sets and let $f_{11}\colon S_1 \to S_1$, $f_{21}\colon S_1 \to S_2$, $f_{32}\colon S_2 \to S_3$, $f_{23}\colon S_3 \to S_2$, and $f_{43}\colon S_3 \to S_4$ be maps. This data then represents a commutative diagram on the directed graph (V, E) where

$$V = \{1, 2, 3, 4\}, \quad E = \{(1, 1), (1, 2), (2, 3), (2, 3), (3, 4)\}.$$

The diagram is represented as

$$\begin{tikzcd} S_1 \arrow[loop left, "f_{11}"] \arrow[r, "f_{21}"] & S_2 \arrow[r, "f_{32}"] & S_3 \arrow[l, bend left, "f_{23}"] \arrow[r, "f_{43}"] & S_4 \end{tikzcd}$$

While it is possible to write down conditions for this diagram to be commutative, there will be infinitely many such conditions. In practice, one encounters commutative diagrams with only finitely many paths with a given source and target. This example, therefore, is not so interesting as a commutative diagram, but is more interesting as a signal flow graph, as we shall see ***missing stuff***. •

Exercises

1.3.1 Let S, T, U, and V be sets, and let $f\colon S \to T$, $g\colon T \to U$, and $h\colon U \to V$ be maps. Show that $h \circ (g \circ f) = (h \circ g) \circ f$.

1.3.2 Let S, T, and U be sets and let $f\colon S \to T$ and $g\colon T \to U$ be maps. Show that $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$ for every subset $C \subseteq U$.

1.3.3 Let S and T be sets, let $f\colon S \to T$, and let $B \subseteq T$. Show that $f^{-1}(T \setminus B) = S \setminus f^{-1}(B)$.

1.3.4 If S, T, and U are sets and if $f\colon S \to T$ and $g\colon T \to U$ are bijections, then show that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

1.3.5 Let S, T and U be sets and let $f\colon S \to T$ and $g\colon T \to U$ be maps.

(a) Show that if f and g are injective, then so too is $g \circ f$.

(b) Show that if f and g are surjective, then so too is $g \circ f$.

1.3.6 Let S and T be sets, let $f\colon S \to T$ be a map, and let $A \subseteq S$ and $B \subseteq T$. Do the following:

(a) show that if f is injective then $A = f^{-1}(f(A))$;

(b) show that if f is surjective then $f(f^{-1}(B)) = B$.

1.3.7 Let S and T be sets and let $f\colon S \to T$ be a map.

(a) Show that if f is invertible as a map, then "the relation of its inverse is the inverse of its relation." (Part of the question is to precisely understand the statement in quotes.)

(b) Show that the inverse of the relation defined by f is itself the relation associated to a function if and only if f is invertible.

1.3.8 Show that equivalence of sets, as in Remark 1.3.7–2, is an "equivalence relation"[5] on collection of all sets.

[5] The quotes are present because the notion of equivalence relation, as we have defined it, applies to sets. However, there is no set containing all sets; see Section **??**.

Section 1.4

Indexed families of sets and general Cartesian products

In this section we discuss general collections of sets, and general collections of members of sets. In Section 1.1.3 we considered Cartesian products of a finite collection of sets. In this section, we wish to extend this to allow for an arbitrary collection of sets. The often used idea of an index set is introduced here, and will come up on many occasions in the text.

Do I need to read this section? The idea of a general family of sets, and notions related to it, do not arise in a lot of places in these volumes. But they do arise. The ideas here are simple (although the notational nuances can be confusing), and so perhaps can be read through. But the reader in a rush can skip the material, knowing they can look back on it if necessary. •

1.4.1 Indexed families and multisets

Recall that when talking about sets, a set is determined only by the concept of membership. Therefore, for example, the sets $\{1,2,2,1,2\}$ and $\{1,2\}$ are the same since they have the same members. However, what if one wants to consider a set with two 1's and three 2's? The way in which one does this is by the use of an index to label the members of the set.

1.4.1 Definition (Indexed family of elements) Let A and S be sets. An ***indexed family of elements*** of S with ***index set*** A is a map $f\colon A \to S$. The element $f(a) \in S$ is sometimes denoted as x_a and the indexed family is denoted as $(x_a)_{a\in A}$. •

missing stuff

With the notion of an indexed family we can make sense of "repeated entries" in a set, as is shown in the first of these examples.

1.4.2 Examples (Indexed family)

1. Consider the two index sets $A_1 = \{1,2,3,4,5\}$ and $A_2 = \{1,2\}$ and let S be the set of natural numbers. Then the functions $f_1\colon A_1 \to S$ and $f_2\colon A_2 \to S$ defined by

$$f_1(1) = 1,\ f_1(2) = 2,\ f_1(3) = 2,\ f_1(4) = 1,\ f_1(5) = 2,$$
$$f_2(1) = 1,\ f_2(2) = 2,$$

   give the indexed families $(x_1 = 1, x_2 = 2, x_3 = 2, x_4 = 1, x_5 = 2)$ and $(x_1 = 1, x_2 = 2)$, respectively. In this way we can arrive at a set with two 1's and three 2's, as desired. Moreover, each of the 1's and 2's is assigned a specific place in the list $(x_1, \dots, x_5)$.

2. Any set S gives rise in a natural way to an indexed family of elements of S indexed by S itself: $(x)_{x\in S}$. •

We can then generalise this notion to an indexed family of sets as follows.

1.4.3 Definition (Indexed family of sets) Let A and S be sets. An ***indexed family of subsets*** of S with ***index set*** A is an indexed family of elements of $\mathbf{2}^S$ with index set A. Thus an indexed family of subsets of S is denoted by $(S_a)_{a\in A}$ where $S_a \subseteq S$ for $a \in A$. •

We use the notation $\cup_{a\in A} S_a$ and $\cap_{a\in A} S_a$ to denote the union and intersection of an indexed family of subsets indexed by A. Similarly, when considering the disjoint union of an indexed family of subsets indexed by A, we define this to be

$$\overset{\circ}{\bigcup_{a\in A}} S_a = \cup_{a\in A}(\{a\} \times S_a).$$

Thus an element in the disjoint union has the form (a, x) where $x \in S_a$. Just as with the disjoint union of a pair of sets, the disjoint union of a family of sets keeps track of the set that element belongs to, now labelled by the index set A, along with the element. A family of sets $(S_a)_{a\in A}$ is ***pairwise disjoint*** if, for every distinct $a_1, a_2 \in A$, $S_{a_1} \cap S_{a_2} = \emptyset$.

Often when one writes $(S_a)_{a\in A}$, one omits saying that the family is "indexed by A," this being understood from the notation. Moreover, many authors will say things like, "Consider the family of sets $\{S_a\}$," so omitting any reference to the index set. In such cases, the index set is usually understood (often it is $\mathbb{Z}_{>0}$). However, we shall not use this notation, and will always give a symbol for the index set.

Sometimes we will simply say something like, "Consider a family of sets $(S_a)_{a\in A}$." When we say this, we tacitly suppose there to be a set S which contains each of the sets S_a as a subset; the union of the sets S_a will serve to give such a set.

There is an alternative way of achieving the objective of allowing sets where the same member appears multiple times.

1.4.4 Definition (Multiset, submultiset) A ***multiset*** is an ordered pair (S, ϕ) where S is a set and $\phi\colon S \to \mathbb{Z}_{\geq 0}$ is a map. A multiset (T, ψ) is a ***submultiset*** of (S, ϕ) if $T \subseteq S$ and if $\psi(x) \leq \phi(x)$ for every $x \in T$. •

This is best illustrated by examples.

1.4.5 Examples (Multisets)

1. The multiset alluded to at the beginning of this section is (S, ϕ) with $S = \{1, 2\}$, and $\phi(1) = 2$ and $\phi(2) = 3$. Note that some information is lost when considering the multiset (S, ϕ) as compared to the indexed family $(1, 2, 2, 1, 2)$; the order of the elements is now immaterial and only the number of occurrences is accounted for.
2. Any set S can be thought of as a multiset (S, ϕ) where $\phi(x) = 1$ for each $x \in S$.
3. Let us give an example of how one might use the notion of a multiset. Let $P \subseteq \mathbb{Z}_{>0}$ be the set of prime numbers and let S be the set $\{2, 3, 4, \dots\}$ of integers greater than 1. As we shall prove in Corollary **??**, every element $n \in S$ can be written in a unique way as $n = p_1^{k_1} \cdots p_m^{k_m}$ for distinct primes $p_1, \dots, p_m$ and for

$k_1, \dots, k_m \in \mathbb{Z}_{>0}$. Therefore, for every $n \in S$ there exists a unique multiset (P, ϕ_n) defined by

$$\phi_n(p) = \begin{cases} k_j, & p = p_j, \\ 0, & \text{otherwise}, \end{cases}$$

understanding that $k_1, \dots, k_m$ and $p_1, \dots, p_m$ satisfy $n = p_1^{k_1} \cdots p_m^{k_m}$. •

1.4.6 Notation (Sets and multisets from indexed families of elements) Let A and S be sets and let $(x_a)_{a \in A}$ be an indexed family of elements of S. If for each $x \in S$ the set $\{a \in A \mid x_a = x\}$ is finite, then one can associate to $(x_a)_{a \in A}$ a multiset (S, ϕ) by

$$\phi(x) = \operatorname{card}\{a \in A \mid x_a = x\}.$$

This multiset is denoted by $\{x_a\}_{a \in A}$. One also has a subset of S associated with the family $(x_a)_{a \in A}$. This is simply the set

$$\{x \in S \mid x = x_a \text{ for some } a \in A\}.$$

This set is denoted by $\{x_a \mid a \in A\}$. Thus we have three potentially quite different objects:

$$(x_a)_{a \in A}, \quad \{x_a\}_{a \in A}, \quad \{x_a \mid a \in A\},$$

arranged in decreasing order of information prescribed (be sure to note that the multiset in the middle is only defined when the sets $\{a \in A \mid x_a = x\}$ are finite). This is possibly confusing, although there is not much in it, really.

For example, the indexed family $(1, 2, 2, 1, 2)$ gives the multiset denoted $\{1, 1, 2, 2, 2\}$ and the set $\{1, 2\}$. Now, this is truly confusing since there is no notational discrimination between the *set* $\{1, 1, 2, 2, 2\}$ (which is simply the set $\{1, 2\}$) and the *multiset* $\{1, 1, 2, 2, 2\}$ (which is not the set $\{1, 2\}$). However, the notation is standard, and the hopefully the intention will be clear from context.

If the map $a \mapsto x_a$ is injective, i.e., the elements in the family $(x_a)_{a \in A}$ are distinct, then the three objects are in natural correspondence with one another. For this reason we can sometimes be a bit lax in using one piece of notation over another. •

1.4.2 General Cartesian products

Before giving general definitions, it pays to revisit the idea of the Cartesian product $S_1 \times S_2$ of sets S_1 and S_2 as defined in Section 1.1.3 (the reason for our change from S and T to S_1 and S_2 will become clear shortly). Let $A = \{1, 2\}$, and let $f \colon A \to S_1 \cup S_2$ be a map satisfying $f(1) \in S_1$ and $f(2) \in S_2$. Then $(f(1), f(2)) \in S_1 \times S_2$. Conversely, given a point $(x_1, x_2) \in S_1 \times S_2$, we define a map $f \colon A \to S_1 \cup S_2$ by $f(1) = x_1$ and $f(2) = x_2$, noting that $f(1) \in S_1$ and $f(2) \in S_2$.

The punchline is that, for a pair of sets S_1 and S_2, their Cartesian product is in 1–1 correspondence with maps f from $A = \{1, 2\}$ to $S_1 \cup S_1$ having the property that $f(x_1) \in S_1$ and $f(x_2) \in S_2$. There are two things to note here: (1) the use of the set A to label the sets S_1 and S_2 and (2) the alternative characterisation of the Cartesian product.

Now we generalise the Cartesian product to families of sets.

1.4.7 Definition (Cartesian product) The ***Cartesian product*** of a family of sets $(S_a)_{a\in A}$ is the set

$$\prod_{a\in A} S_a = \{f\colon A \to \cup_{a\in A} S_a \mid f(a) \in S_a\}. \qquad \bullet$$

Note that the analogue to the ordered pair in a general Cartesian product is simply the set $f(A)$ for some $f \in \prod_{a\in A} S_a$. The reader should convince themselves that this is indeed the appropriate generalisation.

1.4.3 Sequences

The notion of a sequence is very important for us, and we give here a general definition for sequences in arbitrary sets.

1.4.8 Definition (Sequence, subsequence) Let S be a set.

(i) A ***sequence*** in S is an indexed family $(x_j)_{j\in\mathbb{Z}_{>0}}$ of elements of S with index set $\mathbb{Z}_{>0}$.

(ii) A ***subsequence*** of a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ in S is a map $f\colon A \to S$ where

(a) $A \subseteq \mathbb{Z}_{>0}$ is a nonempty set with no upper bound and

(b) $f(k) = x_k$ for all $k \in A$.

If the elements in the set A are ordered as $j_1 < j_2 < j_3 < \cdots$, then the subsequence may be written as $(x_{j_k})_{k\in\mathbb{Z}_{>0}}$. •

Note that in a sequence the location of the elements is important, and so the notation $(x_j)_{j\in\mathbb{Z}_{>0}}$ is the correct choice. It is, however, not uncommon to see sequences denoted $\{x_j\}_{j\in\mathbb{Z}_{>0}}$. According to Notation 1.4.6 this would imply that the same element in S could only appear in the list $(x_j)_{j\in\mathbb{Z}_{>0}}$ a finite number of times. However, this is often not what is intended. However, there is seldom any real confusion induced by this, but the reader should simply be aware that our (not uncommon) notational pedantry is not universally followed.

1.4.4 Directed sets and nets

What we discuss in this section is a generalisation of the notion of a sequence. A sequence is a collection of objects where there is a natural order to the objects inherited from the total order of $\mathbb{Z}_{>0}$.

First we define the index sets for this more general type of sequence.

1.4.9 Definition (Directed set) A ***directed set*** is a partially ordered set $(D, \preceq)$ with the property that, for $x, y \in D$, there exists $z \in D$ such that $x \preceq z$ and $y \preceq z$. •

Thus for any two elements in a directed set D it is possible to find an element greater than either, relative to the specified partial order. Let us give some examples to clarify this.

1.4.10 Examples (Directed sets)

1. The set $(\mathbb{Z}_{>0}, \leq)$ is a directed set since clearly one can find a natural number exceeding any two specified natural numbers.
2. The partially ordered set $([0, \infty), \leq)$ is similarly a directed set.
3. The partially ordered set $((0, 1], \geq)$ is also a directed set since, given $x, y \in (0, 1]$, one can find an element of $(0, 1]$ which is smaller than either x or y.
4. Next take $D = \mathbb{R} \setminus \{x_0\}$ and consider the partial order $\preceq$ on D defined by $x \preceq y$ if $|x - x_0| \leq |y - y_0|$. This may be shown to be a directed set since, given two elements $x, y \in \mathbb{R} \setminus \{x_0\}$, one can find another element of $\mathbb{R} \setminus \{x_0\}$ which is closer to x_0 than either x or y.
5. Let S be a set with more than one element and consider the partially ordered set $(\mathbf{2}^S \setminus \{\emptyset\}, \preceq)$ specified by $A \preceq B$ if $A \supseteq B$. This is readily verified to be a partial order. However, this order does not make $(S, \supseteq)$ a directed set. Indeed, suppose that $A, B \in \mathbf{2}^S \setminus \{\emptyset\}$ are disjoint. Since the only set contained in both A and B is the empty set, it follows that there is no element $T \in \mathbf{2}^S \setminus \{\emptyset\}$ for which $A \supseteq T$ and $B \supseteq T$. •

The next definition is of the generalisation of sequences built on the more general notion of index set given by a directed set.

1.4.11 Definition (Net) Let $(D, \preceq)$ be a directed set. A ***net*** in a set S defined on D is a map $\phi \colon D \to S$ from D into S. •

As with a sequence, it is convenient to instead write $\{x_\alpha\}_{\alpha \in D}$ where $x_\alpha = \phi(\alpha)$ for a net. The idea here is that a net generalises the notion of a sequence to the case where the index set may not be countable and where the order is more general than the total order of $\mathbb{Z}$.

Exercises

1.4.1

Section 1.5

Some words about proving things

Rigour is an important part of the presentation in this series, and if you are so unfortunate as to be using these books as a text, then hopefully you will be asked to prove some things, for example, from the exercises. In this section we say a few (almost uselessly) general things about techniques for proving things. We also say some things about poor proof technique, much (but not all) of which is delivered with tongue in cheek. The fact of the matter is that the best way to become proficient at proving things is to (1) read a lot of (needless to say, good) proofs, and (2) most importantly, get lots of practice. What is certainly true is that it much easier to begin your theorem-proving career by proving simple things. In this respect, the proofs and exercises in this chapter are good ones. Similarly, many of the proofs and exercises in Chapters 4 and **??** provide a good basis for honing one's theorem-proving skills. By contrast, some of the results in Chapter 2 are a little more sophisticated, while still not difficult. As we progress through the preparatory material, we shall increasingly encounter material that is quite challenging, and so proofs that are quite elaborate. The neophyte should not be so ambitious as to tackle these early on in their mathematical development.

Do I need to read this section? Go ahead, read it. It will be fun. •

1.5.1 Legitimate proof techniques

The techniques here are the principle ones use in proving simple results. For very complicated results, many of which appear in this series, one is unlikely to get much help from this list.

1. *Proof by definition:* Show that the desired proposition follows directly from the given definitions and assumptions. Theorems that have already been proven to follow from the definitions and assumptions may also be used. Proofs of this sort are often abbreviated by "This is obvious." While this may well be true, it is better to replace this hopelessly vague assertion with something more meaningful like "This follows directly from the definition."
2. *Proof by contradiction:* Assume that the hypotheses of the desired proposition hold, but that the conclusions are false, and make no other assumption. Show that this leads to an impossible conclusion. This implies that the assumption must be false, meaning the desired proposition is true.
3. *Proof by induction:* In this method one wishes to prove a proposition for an enumerable number of cases, say $1, 2, \dots, n, \dots$. One first proves the proposition for case 1. Then one proves that, if the proposition is true for the nth case, it is true for the $(n + 1)$st case.
4. *Proof by exhaustion:* One proves the desired proposition to be true for all cases. This method only applies when there is a *finite* number of cases.

5. *Proof by contrapositive:* To show that proposition A implies proposition B, one shows that proposition B *not* being true implies that proposition A is *not* true. It is common to see newcomers get proof by contrapositive and proof by contradiction confused.
6. *Proof by counterexample:* This sort of proof is typically useful in showing that some general assertion *does not* hold. That is to say, one wishes to show that a certain conclusion does not follow from certain hypotheses. To show this, it suffices to come up with a single example for which the hypotheses hold, but the conclusion does not. Such an example is called a ***counterexample***.

1.5.2 Improper proof techniques

Many of these seem so simple that a first reaction is, "Who would be dumb enough to do something so obviously incorrect." However, it is easy, and sometimes tempting, to hide one of these incorrect arguments inside something complicated.

1. *Proof by reverse implication:* To prove that A implies B, shows that B implies A.
2. *Proof by half proof:* One is required to show that A and B are equivalent, but one only shows that A implies B. Note that the appearance of "if and only if" means that you have two implications to prove!
3. *Proof by example:* Show only a single case among many. Assume that only a single case is sufficient (when it is not) or suggest that the proof of this case contains most of the ideas of the general proof.
4. *Proof by picture:* A more convincing form of proof by example. Pictures can provide nice illustrations, but suffice in no part of a rigorous argument.
5. *Proof by special methods:* You are allowed to divide by zero, take wrong square roots, manipulate divergent series, etc.
6. *Proof by convergent irrelevancies:* Prove a lot of things related to the desired result.
7. *Proof by semantic shift:* Some standard but inconvenient definitions are changed for the statement of the result.
8. *Proof by limited definition:* Define (or implicitly assume) a set S, for which all of whose elements the desired result is true, then announce that in the future only members of the set S will be considered.
9. *Proof by circular cross-reference:* Delay the proof of a lemma until many theorems have been derived from it. Use one or more of these theorems in the proof of the lemma.
10. *Proof by appeal to intuition:* Cloud-shaped drawings frequently help here.
11. *Proof by elimination of counterexample:* Assume the hypothesis is true. Then show that a counterexample cannot exist. (This is really just a well-disguised proof by reverse implication.) A common variation, known as "begging the question" involves getting deep into the proof and then using a step that assumes the hypothesis.

12. *Proof by obfuscation:* A long plotless sequence of true and/or meaningless syntactically related statements.
13. *Proof by cumbersome notation:* Best done with access to at least four alphabets and special symbols. Can help make proofs by special methods look more convincing.
14. *Proof by cosmology:* The negation of a proposition is unimaginable or meaningless.
15. *Proof by reduction to the wrong problem:* To show that the result is true, compare (reduce/translate) the problem (in)to another problem. This is valid if the other problem is then solvable. The error lies in comparing to an unsolvable problem.

Exercises

1.5.1 Find the flaw in the following inductive "proof" of the fact that, in any class, if one selects a subset of students, they will have received the same grade.

> Suppose that we have a class with students $S = \{S_1, \dots, S_m\}$. We shall prove by induction on the size of the subset that any subset of students receive the same grade. For a subset $\{S_{j_1}\}$, the assertion is clearly true. Now suppose that the assertion holds for all subsets of S with k students with $k \in \{1, \dots, l\}$, and suppose we have a subset $\{S_{j_1}, \dots, S_{j_l}, S_{j_{l+1}}\}$ of $l + 1$ students. By the induction hypothesis, the students from the set $\{S_{j_1}, \dots, S_{j_l}\}$ all receive the same grade. Also by the induction hypothesis, the students from the set $\{S_2, \dots, S_{j_l}, S_{j_{l+1}}\}$ all receive the same grade. In particular, the grade received by student $S_{j_{l+1}}$ is the same as the grade received by student S_{j_l}. But this is the same as the grade received by students $S_{j_1}, \dots, S_{j_{l-1}}$, and so, by induction, we have proved that all students receive the same grade.

In the next exercise you will consider one of Zeno's paradoxes. Zeno[6] is best known for having developed a collection of paradoxes, some of which touch surprisingly deeply on mathematical ideas that were not perhaps fully appreciated until the 19th century. Many of his paradoxes have a flavour similar to the one we give here, which may be the most commonly encountered during dinnertime conversations.

1.5.2 Consider the classical problem of the Achilles chasing the tortoise. A tortoise starts off a race T seconds before Achilles. Achilles, of course, is faster than the tortoise, but we shall argue that, despite this, Achilles will actually never overtake the tortoise.

> At time T when Achilles starts after the tortoise, the tortoise will be some distance d_1 ahead of Achilles. Achilles will reach this point after some time t_1. But, during the time it took Achilles to travel distance d_1, the tortoise will have moved along to some point d_2 ahead of d_1. Achilles will then take a time t_2 to travel the distance

[6] Zeno of Elea (~490BC–~425BC) was an Italian born philosopher of the Greek school.

> d_2. But by then the tortoise will have travelled another distance d_3. This clearly will continue, and when Achilles reaches the point where the tortoise was at some moment before, the tortoise will have moved inexorably ahead. Thus Achilles will never actually catch up to the tortoise.

What is the flaw in the argument?

Chapter 2

Real numbers and their properties

Real numbers and functions of real numbers form an integral part of mathematics. Certainly all students in the sciences receive basic training in these ideas, normally in the form of courses on calculus and differential equations. In this chapter we establish the basic properties of the set of real numbers and of functions defined on this set. In particular, using the construction of the integers in Section **??** as a starting point, we *define* the set of real numbers, thus providing a fairly firm basis on which to develop the main ideas in these volumes. We follow this by discussing various structural properties of the set of real numbers. These cover both algebraic properties (Section 2.2.1) and topological properties (Section 2.5). After this, we discuss important ideas like continuity and differentiability of real-valued functions of a real variable.

Do I need to read this chapter? Yes you do, unless you already know its contents. While the construction of the real numbers in Section 2.1 is perhaps a little bit of an extravagance, it does set the stage for the remainder of the material. Moreover, the material in the remainder of the chapter is, in some ways, the backbone of the mathematical presentation. We say this for two reasons.

1. The technical material concerning the structure of the real numbers is, very simply, assumed knowledge for reading everything else in the series.
2. The *ideas* introduced in this chapter will similarly reappear constantly throughout the volumes in the series. But here, many of these ideas are given their most concrete presentation and, as such, afford the inexperienced reader the opportunity to gain familiarity with useful techniques (e.g., the $\epsilon-\delta$ formalism) in a setting where they presumably possess some degree of comfort. This will be crucial when we discuss more abstract ideas in Chapters **??**, **??**, and **??**, to name a few. •

Contents

2.1 Construction of the real numbers 36

2.1.1 Construction of the rational numbers 36

2.1.2 Construction of the real numbers from the rational numbers 41

2.2 Properties of the set of real numbers 46

2.2.1 Algebraic properties of $\mathbb{R}$ 46

2.2.2 The total order on $\mathbb{R}$ . . . 50
2.2.3 The absolute value function on $\mathbb{R}$ . . . 53
2.2.4 Properties of $\mathbb{Q}$ as a subset of $\mathbb{R}$ . . . 54
2.2.5 The extended real line . . . 58
2.2.6 sup and inf . . . 60
2.2.7 Notes . . . 61
2.3 Sequences in $\mathbb{R}$ . . . 63
2.3.1 Definitions and properties of sequences . . . 63
2.3.2 Some properties equivalent to the completeness of $\mathbb{R}$ . . . 65
2.3.3 Tests for convergence of sequences . . . 68
2.3.4 lim sup and lim inf . . . 69
2.3.5 Multiple sequences . . . 72
2.3.6 Algebraic operations on sequences . . . 74
2.3.7 Convergence using $\mathbb{R}$-nets . . . 75
2.3.8 A first glimpse of Landau symbols . . . 80
2.3.9 Notes . . . 82
2.4 Series in $\mathbb{R}$ . . . 84
2.4.1 Definitions and properties of series . . . 84
2.4.2 Tests for convergence of series . . . 90
2.4.3 e and π . . . 94
2.4.4 Doubly infinite series . . . 98
2.4.5 Multiple series . . . 100
2.4.6 Algebraic operations on series . . . 101
2.4.7 Series with arbitrary index sets . . . 104
2.4.8 Notes . . . 106
2.5 Subsets of $\mathbb{R}$ . . . 110
2.5.1 Open sets, closed sets, and intervals . . . 110
2.5.2 Partitions of intervals . . . 114
2.5.3 Interior, closure, boundary, and related notions . . . 115
2.5.4 Compactness . . . 121
2.5.5 Connectedness . . . 126
2.5.6 Sets of measure zero . . . 126
2.5.7 Cantor sets . . . 130
2.5.8 Notes . . . 132

Section 2.1

Construction of the real numbers

In this section we undertake to define the set of real numbers, using as our starting point the set $\mathbb{Z}$ of integers constructed in Section **??**. The construction begins by building the rational numbers, which are defined, loosely speaking, as fractions of integers. We know from our school days that every real number can be arbitrarily well approximated by a rational number, e.g., using a decimal expansion. We use this intuitive idea as our basis for defining the set of real numbers from the set of rational numbers.

Do I need to read this section? If you feel comfortable with your understanding of what a real number is, then this section is optional reading. However, it is worth noting that in Section 2.1.2 we first use the $\epsilon - \delta$ formalism that is so important in the analysis featured in this series. Readers unfamiliar/uncomfortable with this idea may find this section a good place to get comfortable with this idea. It is also worth mentioning at this point that the $\epsilon - \delta$ formalism is one with which it is difficult to become fully comfortable. Indeed, PhD theses have been written on the topic of how difficult it is for students to fully assimilate this idea. We shall not adopt any unusual pedagogical strategies to address this matter. However, students are well-advised to spend some time understanding $\epsilon - \delta$ language, and instructors are well-advised to appreciate the difficulty students have in coming to grips with it. •

2.1.1 Construction of the rational numbers

The set of rational numbers is, roughly, the set of fractions of integers. However, we do not know what a fraction is. To define the set of rational numbers, we introduce an equivalence relation $\sim$ in $\mathbb{Z} \times \mathbb{Z}_{>0}$ by

$$(j_1, k_1) \sim (j_2, k_2) \quad \iff \quad j_1 \cdot k_2 = j_2 \cdot k_1.$$

We leave to the reader the straightforward verification that this is an equivalence relation. Using this relation we define the rational numbers as follows.

2.1.1 Definition (Rational numbers) A ***rational number*** is an element of $(\mathbb{Z} \times \mathbb{Z}_{>0})/\sim$. The set of rational numbers is denoted by $\mathbb{Q}$. •

2.1.2 Notation (Notation for rationals) For the rational number $[(j,k)]$ we shall typically write $\frac{j}{k}$, reflecting the usual fraction notation. We shall also often write a typical rational number as "q" when we do not care which equivalence class it comes from. We shall denote by 0 and 1 the rational numbers $[(0,1)]$ and $[(1,1)]$, respectively •

The set of rational numbers has many of the properties of integers. For example, one can define addition and multiplication for rational numbers, as well as a total order in the set of rationals. However, there is an important construction that can be made for rational numbers that cannot generally be made for integers, namely that of division. Let us see how this is done.

2.1.3 Definition (Addition, multiplication, and division in $\mathbb{Q}$) Define the operations of ***addition***, ***multiplication***, and ***division*** in $\mathbb{Q}$ by

(i) $[(j_1,k_1)] + [(j_2,k_2)] = [(j_1 \cdot k_2 + j_2 \cdot k_1, k_1 \cdot k_2)]$,

(ii) $[(j_1,k_1)] \cdot [(j_2,k_2)] = [(j_1 \cdot j_2, k_1 \cdot k_2)]$, and

(iii) $[(j_1,k_1)]/[(j_2,k_2)] = [(j_1 \cdot k_2, k_1 \cdot j_2)]$ (we will also write $\frac{[(j_1,k_1)]}{[(j_2,k_2)]}$ for $[(j_1,k_1)]/[(j_2,k_2)]$),

respectively, where $[(j_1,k_1)], [(j_2,k_2)] \in \mathbb{Q}$ and where, in the definition of division, we require that $j_2 \neq 0$. We will sometimes omit the "$\cdot$" when in multiplication. •

We leave to the reader as Exercise 2.1.1 the straightforward task of showing that these definitions are independent of choice of representatives in $\mathbb{Z} \times \mathbb{Z}_{>0}$. We also leave to the reader the assertion that, with respect to Notation 2.1.2, the operations of addition, multiplication, and division of rational numbers assume the familiar form:

$$\frac{j_1}{k_1} + \frac{j_2}{k_2} = \frac{j_1 \cdot k_2 + j_2 \cdot k_1}{k_1 \cdot k_2}, \quad \frac{j_1}{k_1} \cdot \frac{j_2}{k_2} = \frac{j_1 \cdot j_2}{k_2 \cdot k_2}, \quad \frac{\frac{j_1}{k_1}}{\frac{j_2}{k_2}} = \frac{j_1 \cdot k_2}{k_1 \cdot j_2}.$$

For the operation of division, it is convenient to introduce a new concept. Given $[(j,k)] \in \mathbb{Q}$ with $j \neq 0$, we define $[(j,k)]^{-1} \in \mathbb{Q}$ by $[(k,j)]$. With this notation, division then can be written as $[(j_1,k_1)]/[(j_2,k_2)] = [(j_1,k_1)] \cdot [(j_2,k_2)]^{-1}$. Thus division is really just multiplication, as we already knew. Also, if $q \in \mathbb{Q}$ and if $k \in \mathbb{Z}_{\geq 0}$, then we define $q^k \in \mathbb{Q}$ inductively by $q^0 = 1$ and $q^{k^+} = q^k \cdot q$. The rational number q^k is the kth ***power*** of q.

Let us verify that the operations above satisfy the expected properties. Note that there are now some new properties, since we have the operation of division, or multiplicative inversion, to account for. As we did for integers, we shall write $-q$ for $-1 \cdot q$.

2.1.4 Proposition (Properties of addition and multiplication in $\mathbb{Q}$) *Addition and multiplication in $\mathbb{Q}$ satisfy the following rules:*

(i) $q_1 + q_2 = q_2 + q_1$, $q_1, q_2 \in \mathbb{Q}$ *(**commutativity** of addition);*

(ii) $(q_1 + q_2) + q_3 = q_1 + (q_2 + q_3)$, $q_1, q_2, q_3 \in \mathbb{Q}$ *(**associativity** of addition);*

(iii) $q + 0 = q$, $q \in \mathbb{Q}$ *(**additive identity**);*

(iv) $q + (-q) = 0$, $q \in \mathbb{Q}$ *(**additive inverse**);*

(v) $q_1 \cdot q_2 = q_2 \cdot q_1$, $q_1, q_2 \in \mathbb{Q}$ *(**commutativity** of multiplication);*

(vi) $(q_1 \cdot q_2) \cdot q_3 = q_1 \cdot (q_2 \cdot q_3)$, $q_1, q_2, q_3 \in \mathbb{Q}$ *(**associativity** of multiplication);*

(vii) $q \cdot 1 = q$, $q \in \mathbb{Q}$ *(**multiplicative identity**);*

(viii) $q \cdot q^{-1} = 1$, $q \in \mathbb{Q} \setminus \{0\}$ *(**multiplicative inverse**);*

(ix) $r \cdot (q_1 + q_2) = r \cdot q_1 + r \cdot q_2$, $r, q_1, q_2 \in \mathbb{Q}$ *(**distributivity**);*

(x) $q^{k_1} \cdot q^{k_2} = q^{k_1+k_2}$, $q \in \mathbb{Q}$, $k_1, k_2 \in \mathbb{Z}_{\geq 0}$.

Moreover, if we define $\mathrm{i}_{\mathbb{Z}}\colon \mathbb{Z} \to \mathbb{Q}$ *by* $\mathrm{i}_{\mathbb{Z}}(\mathrm{k}) = [(\mathrm{k}, 1)]$*, then addition and multiplication in* $\mathbb{Q}$ *agrees with that in* $\mathbb{Z}$*:*

$$\mathrm{i}_{\mathbb{Z}}(\mathrm{k}_1) + \mathrm{i}_{\mathbb{Z}}(\mathrm{k}_2) = \mathrm{i}_{\mathbb{Z}}(\mathrm{k}_1 + \mathrm{k}_2), \quad \mathrm{i}_{\mathbb{Z}}(\mathrm{k}_1) \cdot \mathrm{i}_{\mathbb{Z}}(\mathrm{k}_2) = \mathrm{i}_{\mathbb{Z}}(\mathrm{k}_1 \cdot \mathrm{k}_2).$$

Proof All of these properties follow directly from the definitions of addition and multiplication, using Proposition **??**. ■

Just as we can naturally think of $\mathbb{Z}_{\geq 0}$ as being a subset of $\mathbb{Z}$, so too can we think of $\mathbb{Z}$ as a subset of $\mathbb{Q}$. Moreover, we shall very often do so without making explicit reference to the map $i_{\mathbb{Z}}$.

Next we consider on $\mathbb{Q}$ the extension of the partial order $\leq$ and the strict partial order $<$.

2.1.5 Proposition (Order on $\mathbb{Q}$) *On* $\mathbb{Q}$ *define two relations* $<$ *and* $\leq$ *by*

$$\begin{aligned} [(\mathrm{j}_1, \mathrm{k}_1)] < [(\mathrm{j}_2, \mathrm{k}_2)] \quad &\iff \quad \mathrm{j}_1 \cdot \mathrm{k}_2 < \mathrm{k}_1 \cdot \mathrm{j}_2, \\ [(\mathrm{j}_1, \mathrm{k}_1)] \leq [(\mathrm{j}_2, \mathrm{k}_2)] \quad &\iff \quad \mathrm{j}_1 \cdot \mathrm{k}_2 \leq \mathrm{k}_1 \cdot \mathrm{j}_2. \end{aligned}$$

Then $\leq$ *is a total order and* $<$ *is the corresponding strict partial order.*

Proof First let us show that the relations defined make sense, in that they are independent of choice of representative. Thus we suppose that $[(j_1, k_1)] = [(\tilde{j}_1, \tilde{k}_1)]$ and that $[(j_2, k_2)] = [(\tilde{j}_2, \tilde{k}_2)]$. Then

$$\begin{aligned} & [(j_1, k_1)] \leq [(j_2, k_2)] \\ \iff \quad & j_1 \cdot k_2 \leq k_1 \cdot j_2 \\ \iff \quad & j_1 \cdot k_2 \cdot j_2 \cdot \tilde{k}_2 \cdot \tilde{j}_1 \cdot k_1 \leq k_1 \cdot j_2 \cdot \tilde{j}_2 \cdot k_1 \cdot j_1 \cdot \tilde{k}_1 \\ \iff \quad & (\tilde{j}_1 \cdot \tilde{k}_2) \cdot (j_1 \cdot j_2 \cdot k_1 \cdot k_2) \leq (\tilde{j}_2 \cdot \tilde{k}_1) \cdot (j_1 \cdot j_2 \cdot k_1 \cdot k_2) \\ \iff \quad & \tilde{j}_1 \cdot \tilde{k}_2 \leq \tilde{j}_2 \cdot \tilde{k}_1. \end{aligned}$$

This shows that the definition of $\leq$ is independent of representative. Of course, a similar argument holds for $<$.

That $\leq$ is a partial order, and that $<$ is its corresponding strict partial order, follow from a straightforward checking of the definitions, so we leave this to the reader.

Thus we only need to check that $\leq$ is a total order. Let $[(j_1, k_1)], [(j_2, k_2)] \in \mathbb{Q}$. Then, by the Trichotomy Law for $\mathbb{Z}$, either $j_1 \cdot k_2 < k_1 \cdot j_2$, $k_1 \cdot j_2 < j_1 \cdot k_2$, or $j_1 \cdot k_2 = k_1 \cdot j_2$. But this directly implies that either $[(j_1, k_1)] < [(j_2, k_2)]$, $[(j_2, k_2)] < [(j_1, k_1)]$, or $[(j_1, k_1)] = [(j_2, k_2)]$, respectively. ■

The total order on $\mathbb{Q}$ allows a classification of rational numbers as follows.

2.1.6 Definition (Positive and negative rational numbers) A rational number $q \in \mathbb{Q}$ is:

(i) ***positive*** if $0 < q$;

(ii) ***negative*** if $q < 0$;

(iii) ***nonnegative*** if $0 \leq q$;

(iv) ***nonpositive*** if $q \leq 0$.

The set of positive rational numbers is denoted by $\mathbb{Q}_{>0}$ and the set of nonnegative rational numbers is denoted by $\mathbb{Q}_{\geq 0}$. •

As we did with natural numbers and integers, we isolate the Trichotomy Law.

2.1.7 Corollary (Trichotomy Law for $\mathbb{Q}$) *For* $\mathrm{q}, \mathrm{r} \in \mathbb{Q}$*, exactly one of the following possibilities holds:*

(i) $\mathrm{q} < \mathrm{r}$*;*

(ii) $\mathrm{r} < \mathrm{q}$*;*

(iii) $\mathrm{q} = \mathrm{r}$*.*

The following result records the relationship between the order on $\mathbb{Q}$ and the arithmetic operations.

2.1.8 Proposition (Relation between addition and multiplication and <) *For* $\mathrm{q}, \mathrm{r}, \mathrm{s} \in \mathbb{Q}$*, the following statements hold:*

(i) if $\mathrm{q} < \mathrm{r}$ *then* $\mathrm{q} + \mathrm{s} < \mathrm{r} + \mathrm{s}$*;*

(ii) if $\mathrm{q} < \mathrm{r}$ *and if* $\mathrm{s} > 0$ *then* $\mathrm{s} \cdot \mathrm{q} < \mathrm{s} \cdot \mathrm{r}$*;*

(iii) if $\mathrm{q} < \mathrm{r}$ *and if* $\mathrm{s} < 0$ *then* $\mathrm{s} \cdot \mathrm{r} < \mathrm{s} \cdot \mathrm{q}$*;*

(iv) if $0 < \mathrm{q}, \mathrm{r}$ *then* $0 < \mathrm{q} \cdot \mathrm{r}$*;*

(v) if $\mathrm{q} < \mathrm{r}$ *and if either*

(a) $0 < \mathrm{q}, \mathrm{r}$ *or*

(b) $\mathrm{q}, \mathrm{r} < 0$*,*

then $\mathrm{r}^{-1} < \mathrm{q}^{-1}$*.*

Proof (i) Write $q = [(j_q, k_q)]$, $r = [(j_r, k_r)]$, and $s = [(j_s, k_s)]$. Since $q < r$, $j_q \cdot k_r \leq j_r \cdot k_q$. Therefore,

$$\begin{aligned} & j_q \cdot k_r \cdot k_s^2 < j_r \cdot k_q \cdot k_s^2 \\ \Longrightarrow \quad & j_q \cdot k_r \cdot k_s^2 + j_s \cdot k_q \cdot k_r \cdot k_s < j_r \cdot k_q \cdot k_s^2 + j_2 \cdot k_q \cdot k_r \cdot k_s, \end{aligned}$$

using Proposition **??**. This last inequality is easily seen to be equivalent to $q + s < r + s$.

(ii) Write $q = [(j_q, k_q)]$, $r = [(j_r, k_r)]$, and $s = [(j_s, k_s)]$. Since $s > 0$ it follows that $j_s > 0$. Since $q \leq r$ it follows that $j_q \cdot k_r \leq j_r \cdot k_q$. From Proposition **??** we then have

$$j_q \cdot j_s \cdot j_s \cdot k_s \leq j_r \cdot k_q \cdot j_s \cdot k_s,$$

which is equivalent to $s \cdot q \leq s \cdot r$ by definition of multiplication.

(iii) The result here follows, as does (ii), from Proposition **??**, but now using the fact that $j_s < 0$.

(iv) This is a straightforward application of the definition of multiplication and <.

(v) This follows directly from the definition of <. ■

The final piece of structure we discuss for rational numbers is the extension of the absolute value function defined for integers.

2.1.9 Definition (Rational absolute value function) The ***absolute value function*** on $\mathbb{Q}$ is the map from $\mathbb{Q}$ to $\mathbb{Q}_{\geq 0}$, denoted by $q \mapsto |q|$, defined by

$$|q| = \begin{cases} q, & 0 < q, \\ 0, & q = 0, \\ -q, & q < 0. \end{cases} \qquad \bullet$$

The absolute value function on $\mathbb{Q}$ has properties like that on $\mathbb{Z}$.

2.1.10 Proposition (Properties of absolute value on $\mathbb{Q}$) *The following statements hold:*

(i) $|q| \geq 0$ *for all* $q \in \mathbb{Q}$;

(ii) $|q| = 0$ *if and only if* $q = 0$;

(iii) $|r \cdot q| = |r| \cdot |q|$ *for all* $r, q \in \mathbb{Q}$;

(iv) $|r + q| \leq |r| + |q|$ *for all* $r, q \in \mathbb{Q}$ *(**triangle inequality**)*;

(v) $|q^{-1}| = |q|^{-1}$ *for all* $q \in \mathbb{Q} \setminus \{0\}$.

Proof Parts (i), (ii), and (v), follow directly from the definition, and part (iii) follows in the same manner as the analogous statement in Proposition **??**. Thus we have only to prove part (iv). We consider various cases.

1. $|r| \leq |q|$:
   (a) $0 \geq r, q$: Since $|r + q| = r + q$, and $|r| = r$ and $|q| = q$, this follows directly.
   (b) $r < 0$, $0 \leq q$: Let $r = [(j_r, k_r)]$ and $q = [(j_q, k_q)]$. Then $r < 0$ gives $j_r < 0$ and $0 \leq q$ gives $j_q \geq 0$. We now have

$$|r + q| = \left|\frac{j_r \cdot k_q + j_q \cdot k_r}{k_r \cdot k_q}\right| = \frac{|j_r \cdot k_q + j_q \cdot k_r|}{k_r \cdot k_q}$$

   and

$$|r| + |q| = \frac{|j_r| \cdot k_q + |j_q| \cdot k_r}{k_r \cdot k_q}.$$

   Therefore,

$$\begin{aligned} |r + q| &= \frac{|j_r \cdot k_q + j_q \cdot k_r|}{k_r \cdot k_q} \\ &\leq \frac{|j_r| \cdot k_q + |j_q| \cdot k_r}{k_r \cdot k_q} \\ &= |r| + |q|, \end{aligned}$$

   where we have used Proposition 2.1.8.
   (c) $r, q < 0$: Here $|r + q| = |-r + (-q)| = |-(r + q)| = -(r + q)$, and $|r| = -r$ and $|q| = -q$, so the result follows immediately.
2. $|q| \leq |r|$: This argument is the same as above, swapping r and q. ■

2.1.11 Remark Having been quite fussy about how we arrived at the set of integers and the set of rational numbers, and about characterising their important properties, we shall now use standard facts about these, some of which we may not have proved, but which can easily be proved using the definitions of $\mathbb{Z}$ and $\mathbb{Q}$. Some of the arithmetic properties of $\mathbb{Z}$ and $\mathbb{Q}$ that we use without comment are in fact proved in Section **??** in the more general setting of rings. However, we anticipate that most readers will not balk at the instances where we use unproved properties of integers and rational numbers. •

2.1.2 Construction of the real numbers from the rational numbers

Now we use the rational numbers as the building block for the real numbers. The idea of this construction, which was originally due to Cauchy[1], is the intuitive idea that the rational numbers may be used to approximate well a real number. For example, we learn in school that any real number is expressible as a decimal expansion (see Exercise 2.4.8 for the precise construction of a decimal expansion). However, any finite length decimal expansion (and even some infinite length decimal expansions) is a rational number. So one could *define* real numbers as a limit of decimal expansions in some way. The problem is that there may be multiple decimal expansions giving rise to the same real number. For example, the decimal expansions 1.0000 and 0.9999... represent the same real number. The way one gets around this potential problem is to use equivalence classes, of course. But equivalence classes of what? This is where we begin the presentation, proper.

2.1.12 Definition (Cauchy sequence, convergent sequence) Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{Q}$. The sequence:

(i) is a ***Cauchy sequence*** if, for each $\epsilon \in \mathbb{Q}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_k| < \epsilon$ for $j, k \geq N$;

(ii) ***converges to*** $\mathbf{q_0}$ if, for each $\epsilon \in \mathbb{Q}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_0| < \epsilon$ for $j \geq N$.

(iii) is ***bounded*** if there exists $M \in \mathbb{Q}_{>0}$ such that $|q_j| < M$ for each $j \in \mathbb{Z}_{>0}$. •

The set of Cauchy sequences in $\mathbb{Q}$ is denoted by $\mathrm{CS}(\mathbb{Q})$. A sequence converging to q_0 has q_0 as its ***limit***. •

The idea of a Cauchy sequence is that the terms in the sequence can be made arbitrarily close as we get to the tail of the sequence. A convergent sequence, however, gets closer and closer to its limit as we get to the tail of the sequence. Our instinct is probably that there is a relationship between these two ideas. One thing that is true is the following.

2.1.13 Proposition (Convergent sequences are Cauchy) *If a sequence* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* $\mathrm{q_0}$*, then it is a Cauchy sequence.*

Proof Let $\epsilon \in \mathbb{Q}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_0| < \frac{\epsilon}{2}$ for $j \geq N$. Then, for $j, k \geq N$ we have

$$|q_j - q_k| = |q_j - q_0 - q_k + q_0| = |q_j - q_0| + |q_k - q_0| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,$$

using the triangle inequality of Proposition 2.1.10. ■

Cauchy sequences have the property of being bounded.

[1] The French mathematician Augustin Louis Cauchy (1789–1857) worked in the areas of complex function theory, partial differential equations, and analysis. His collected works span twenty-seven volumes.

2.1.14 Proposition (Cauchy sequences are bounded) *If* $(q_j)_{j\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence, then it is bounded.*

Proof Choose $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_k| < 1$ for $j,k \in \mathbb{Z}_{>0}$. Then take M_N to be the largest of the nonnegative rational numbers $|q_1|,\dots,|q_N|$. Then, for $j \geq N$ we have, using the triangle inequality,

$$|q_j| = |q_j - q_N + q_N| \leq |q_j - q_N| + |q_N| < 1 + M_N,$$

giving the result by taking $M = M_N + 1$. ■

The question as to whether there are nonconvergent Cauchy sequences is now the obvious one.

2.1.15 Example (Nonconvergent Cauchy sequences in $\mathbb{Q}$ exist) If one already knows the real numbers exist, it is somewhat easy to come up with Cauchy sequences in $\mathbb{Q}$. However, to fabricate one "out of thin air" is not so easy.

For $k \in \mathbb{Z}_{>0}$, since $2k + 5 > k + 4$, it follows that $2^{2k+5} - 2^{k+4} > 0$. Let m_k be the smallest nonnegative integer for which

$$m_k^2 \geq 2^{2k+5} - 2^{k+4}. \tag{2.1}$$

The following contains a useful property of m_k.

1 Lemma $m_k^2 \leq 2^{2k+5}$.

Proof First we show that $m_k \leq 2^{k+3}$. Suppose that $m_k > 2^{k+3}$. Then

$$(m_k - 1)^2 > (2^{k+3} - 1)^2 = 2^{2k+6} - 2^{k+4} + 1 = 2(2^{2k+5} - 2^{k+4}) + 1) > 2^{2k+5} - 2^{k+4},$$

which contradicts the definition of m_k.

Now suppose that $m_k^2 > 2^{2k+5}$. Then

$$(m_k - 1)^2 = m_k^2 - 2m_k + 1 > 2^{2k+5} - 2^{k+4} + 1 > 2^{2k+5} - 2^{k+4},$$

again contradicting the definition of m_k. ▼

Now define $q_k = \frac{m_k}{2^{k+2}}$.

2 Lemma $(q_k)_{k\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence.*

Proof By Lemma 1 we have

$$q_k^2 = \frac{m_k^2}{2^{2k+4}} \leq \frac{2^{2k+5}}{2^{2k+4}} = 2, \qquad k \in \mathbb{Z}_{>0},$$

and by (2.1) we have

$$q_k^2 = \frac{m_k^2}{2^{2k+4}} \geq \frac{2^{2k+5}}{2^{2k+4}} - \frac{2^{k+4}}{2^{2k+4}} = 2 - \frac{1}{2k}, \qquad k \in \mathbb{Z}_{>0}.$$

Summarising, we have

$$2-\frac{1}{2^k}\le q_k^2\le 2,\qquad k\in\mathbb{Z}_{>0}. \tag{2.2}$$

Then, for $j,k\in\mathbb{Z}_{>0}$ we have

$$2-\frac{1}{2^k}\le q_k^2\le 2,\ 2-\frac{1}{2^j}\le q_k^2\le 2\quad\Longrightarrow\quad -\frac{1}{2^j}\le q_j^2-q_k^2\le\frac{1}{2^k}.$$

Next we have, from (2.1),

$$q_k^2=\frac{m_k^2}{2^{2k+4}}\ge\frac{2^{2k+5}}{2^{2k+4}}-\frac{2^{k+4}}{2^{2k+4}}=2-\frac{1}{2^k},\qquad k\in\mathbb{Z}_{>0},$$

from which we deduce that $q_k^2\ge 1$, which itself implies that $q_k\ge 1$. Next, using this fact and $(q_j-q_k)^2=(q_j+q_k)(q_j-q_k)$ we have

$$-\frac{1}{2^j}\frac{1}{q_j+q_k}\le q_j-q_k\le\frac{1}{2^j}\frac{1}{q_j+q_k}\quad\Longrightarrow\quad -\frac{1}{2^{j+1}}\le q_j-q_k\le\frac{1}{2^{k+1}},\qquad j,k\in\mathbb{Z}_{>0}. \tag{2.3}$$

Now let $\epsilon\in\mathbb{Q}_{>0}$ and choose $N\in\mathbb{Z}_{>0}$ such that $\frac{1}{2^{N+1}}<\epsilon$. Then we immediately have $|q_j-q_k|<\epsilon$, $j,k\ge N$, using (2.3). ▼

The following result gives the character of the limit of the sequence $(q_k)_{k\in\mathbb{Z}_{>0}}$, were it to be convergent.

3 Lemma *If* q_0 *is the limit for the sequence* $(\mathrm{q}_k)_{k\in\mathbb{Z}_{>0}}$*, then* $\mathrm{q}_0^2=2$.

Proof We claim that if $(q_k)_{k\in\mathbb{Z}_{>0}}$ converges to q_0, then $(q_k^2)_{k\in\mathbb{Z}_{>0}}$ converges to q_0^2. Let $M\in\mathbb{Q}_{>0}$ satisfy $|q_k|<M$ for all $k\in\mathbb{Z}_{>0}$, this being possible by Proposition 2.1.14. Now let $\epsilon\in\mathbb{Q}_{>0}$ and take $N\in\mathbb{Z}_{>0}$ such that

$$|q_k-q_0|<\frac{\epsilon}{M+|q_0|}.$$

Then

$$|q_k^2-q_0^2|=|q_k-q_0||q_k+q_0|<\epsilon,$$

giving our claim.

Finally, we prove the lemma by proving that $(q_k^2)_{k\in\mathbb{Z}_{>0}}$ converges to 2. Indeed, let $\epsilon\in\mathbb{Q}_{>0}$ and note that, if $N\in\mathbb{Z}_{>0}$ is chosen to satisfy $\frac{1}{2^N}<\epsilon$. Then, using (2.2), we have

$$|q_k^2-2|\le\frac{1}{2^k}<\epsilon,\qquad k\ge N,$$

as desired. ▼

Finally, we have the following result, which is contained in the mathematical works of Euclid.

4 **Lemma** *There exists no* $\mathrm{q}_0 \in \mathbb{Q}$ *such that* $\mathrm{q}_0^2 = 2$*.*

Proof Suppose that $q_0^2 = [(j_0, k_0)]$ and further suppose that there is no integer m such that $q_0 = [(mj_0, mk_0)]$. We then have

$$q_0^2 = \frac{j_0^2}{k_0^2} = 2 \qquad \Longrightarrow \qquad j_0^2 = 2k_0^2.$$

Thus j_0^2 is even, and then so too is j_0 (why?). Therefore, $j_0 = 2\tilde{j}_0$ and so

$$q_0^2 = \frac{4\tilde{j}_0^2}{k_0^2} = 2 \qquad \Longrightarrow \qquad k_0^2 = 2\tilde{j}_0^2$$

which implies that k_0^2, and hence k_0 is also even. This contradicts our assumption that there is no integer m such that $q_0 = [(mj_0, mk_0)]$. ▼

With these steps, we have constructed a Cauchy sequence that does not converge. •

Having shown that there are Cauchy sequences that do not converge, the idea is now to define a real number to be, essentially, that to which a nonconvergent Cauchy sequence would converge if only it could. First we need to allow for the possibility, realised in practice, that different Cauchy sequences may converge to the same limit.

2.1.16 Definition (Equivalent Cauchy sequences) Two sequences $(q_j)_{j\in\mathbb{Z}_{>0}}, (r_j)_{j\in\mathbb{Q}} \in \mathrm{CS}(\mathbb{Q})$ are ***equivalent*** if the sequence $(q_j - r_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero. We write $(q_j)_{j\in\mathbb{Z}_{>0}} \sim (r_j)_{j\in\mathbb{Z}_{>0}}$ if the two sequences are equivalent. •

We should verify that this notion of equivalence of Cauchy sequences is indeed an equivalence relation.

2.1.17 Lemma *The relation* $\sim$ *defined in* $\mathrm{CS}(\mathbb{Q})$ *is an equivalence relation.*

Proof It is clear that the relation $\sim$ is reflexive and symmetric. To prove transitivity, suppose that $(q_j)_{j\in\mathbb{Z}_{>0}} \sim (r_j)_{j\in\mathbb{Z}_{>0}}$ and that $(r_j)_{j\in\mathbb{Z}_{>0}} \sim (s_j)_{j\in\mathbb{Z}_{>0}}$. For $\epsilon \in \mathbb{Q}_{>0}$ let $N \in \mathbb{Z}_{>0}$ satisfy

$$|q_j - r_j| < \tfrac{\epsilon}{2}, \ |r_j - s_j| < \tfrac{\epsilon}{2}, \qquad j \geq N.$$

Then, using the triangle inequality,

$$|q_j - s_j| = |q_j - r_j + r_j - s_j| \leq |q_j - r_j| + |r_j - s_j| < \epsilon, \qquad j \geq \mathbb{Z}_{>0},$$

showing that $(q_j)_{j\in\mathbb{Z}_{>0}} \sim (s_j)_{j\in\mathbb{Z}_{>0}}$. ■

We are now prepared to define the set of real numbers.

2.1.18 Definition (Real numbers) A ***real number*** is an element of $\mathrm{CS}(\mathbb{Q})/\sim$. The set of real numbers is denoted by $\mathbb{R}$. •

The definition encodes, in a precise way, our intuition about what a real number is. In the next section we shall examine some of the properties of the set $\mathbb{R}$.

Let us give the notation we will use for real numbers, since clearly we do not wish to write these explicitly as equivalence classes of Cauchy sequences.

2.1.19 Notation (Notation for reals) We shall frequently write a typical element in $\mathbb{R}$ as "x". We shall denote by 0 and 1 the real numbers associated with the Cauchy sequences $(0)_{j\in\mathbb{Z}_{>0}}$ and $(1)_{j\in\mathbb{Z}_{>0}}$. •

Exercises

2.1.1 Show that the definitions of addition, multiplication, and division of rational numbers in Definition 2.1.3 are independent of representative.

2.1.2 Show that the order and absolute value on $\mathbb{Q}$ agree with those on $\mathbb{Z}$. That is to say, show the following:

(a) for $j, k \in \mathbb{Z}$, $j < k$ if and only if $i_{\mathbb{Z}}(j) < i_{\mathbb{Z}}(k)$;

(b) for $k \in \mathbb{Z}$, $|k| = |i_{\mathbb{Z}}(k)|$.

(Note that we see clearly here the abuse of notation that follows from using $<$ for both the order on $\mathbb{Z}$ and $\mathbb{Q}$ and from using $|\cdot|$ as the absolute value both on $\mathbb{Z}$ and $\mathbb{Q}$. It is expected that the reader can understand where the notational abuse occurs.)

2.1.3 Show that the set of rational numbers is countable using an argument along the following lines.

1. Construct a doubly infinite grid in the plane with a point at each integer coordinate. Note that every rational number $q = \frac{n}{m}$ is represented by the grid point (n, m).
2. Start at the "centre" of the grid with the rational number 0 being assigned to the grid point $(0, 0)$, and construct a spiral which passes through each grid point. Note that this spiral should hit every grid point exactly once.
3. Use this spiral to infer the existence of a bijection from $\mathbb{Q}$ to $\mathbb{Z}_{>0}$.

The following exercise leads you through Cantor's famous "diagonal argument" for showing that the set of real numbers is uncountable.

2.1.4 Fill in the gaps in the following construction, justifying all steps.

1. Let $\{x_j \mid j \in \mathbb{Z}_{>0}\}$ be a countable subset of $(0, 1)$.
2. Construct a doubly infinite table for which the kth column of the jth row contains the kth term in the decimal expansion for x_j.
3. Construct $\bar{x} \in (0, 1)$ by declaring the kth term in the decimal expansion for $\bar{x}$ to be different from the kth term in the decimal expansion for x_k.
4. Show that $\bar{x}$ is not an element of the set $\{x_j \mid j \in \mathbb{Z}_{>0}\}$.
   Hint: *Be careful to understand that a real number might have different decimal expansions.*

2.1.5 Show that for any $x \in \mathbb{R}$ and $\epsilon \in \mathbb{R}_{>0}$ there exists $k \in \mathbb{Z}_{>0}$ and an odd integer j such that $|x - \frac{j}{2^k}| < \epsilon$.

Section 2.2

Properties of the set of real numbers

In this section we present some of the well known properties as the real numbers, both algebraic and (referring ahead to the language of Chapter **??**) topological.

Do I need to read this section? Many of the properties given in Sections 2.2.1, 2.2.2 and 2.2.3 will be well known to any student with a high school education. However, these may be of value as a starting point in understanding some of the abstract material in Chapters 4 and **??**. Similarly, the material in Section 2.2.4 is "obvious." However, since this material will be assumed knowledge, it might be best for the reader to at least skim the section, to make sure there is nothing new in it for them. •

2.2.1 Algebraic properties of $\mathbb{R}$

In this section we define addition, multiplication, order, and absolute value for $\mathbb{R}$, mirroring the presentation for $\mathbb{Q}$ in Section 2.1.1. Here, however, the definitions and verifications are not just trivialities, as they are for $\mathbb{Q}$.

First we define addition and multiplication. We do this by defining these operations first on elements of $\mathrm{CS}(\mathbb{Q})$, and then showing that the operations depend only on equivalence class. The following is the key step in doing this.

2.2.1 Proposition (Addition, multiplication, and division of Cauchy sequences) *Let* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \mathrm{CS}(\mathbb{Q})$*. Then the following statements hold.*

(i) The sequence $(\mathrm{q_j}+\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence which we denote by* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}+(\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$*.*

(ii) The sequence $(\mathrm{q_j}\cdot\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence which we denote by* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}\cdot(\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$*.*

(iii) If, for all $\mathrm{j}\in\mathbb{Z}_{>0}$*,* $\mathrm{q_j}\neq 0$ *and if the sequence* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *does not converge to 0, then* $(\mathrm{q_j^{-1}})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence.*

Furthermore, if $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \mathrm{CS}(\mathbb{Q})$ *satisfy*

$$(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, \quad (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}},$$

then

(iv) $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} + (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} = (\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}} + (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$*,*

(v) $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \cdot (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} = (\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \cdot (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$*, and*

(vi) if, for all $\mathrm{j}\in\mathbb{Z}_{>0}$*,* $\mathrm{q_j}, \tilde{\mathrm{q}}_\mathrm{j} \neq 0$ *and if the sequences* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *do not converge to 0, then* $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$*.*

Proof (i) Let $\epsilon \in \mathbb{Q}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ have the property that $|q_j - q_k|, |r_j - r_k| < \frac{\epsilon}{2}$ for all $j, k \geq N$. Then, using the triangle inequality,

$$|(q_j + r_j) - (q_k + r_k)| \leq |q_j - q_k| + |r_j - r_k| = \epsilon, \qquad j, k \geq N.$$

(ii) Let $M \in \mathbb{Q}_{>0}$ have the property that $|q_j|, |r_j| < M$ for all $j \in \mathbb{Z}_{>0}$. For $\epsilon \in \mathbb{Q}_{>0}$ let $N \in \mathbb{Z}_{>0}$ have the property that $|q_j - q_k|, |r_j - r_k| < \frac{\epsilon}{2M}$ for all $j, k \geq N$. Then, using the triangle inequality,

$$\begin{aligned}|(q_j \cdot r_j) - (q_k \cdot r_k)| &= |q_j(r_j - r_k) - r_k(q_k - q_j)| \\ &\leq |q_j||r_j - r_k| + |r_k||q_k - q_j| < \epsilon, \qquad j, k \geq N.\end{aligned}$$

(iii) We claim that if $(q_j)_{j\in\mathbb{Z}_{>0}}$ satisfies the conditions stated, then there exists $\delta \in \mathbb{Q}_{>0}$ such that $|q_k| \geq \delta$ for all $k \in \mathbb{Z}_{>0}$. Indeed, since $(q_j)_{j\in\mathbb{Z}_{>0}}$ does not converge to zero, choose $\epsilon \in \mathbb{Q}_{>0}$ such that, for all $N \in \mathbb{Z}_{>0}$, there exists $j \geq N$ for which $|q_j| \geq \epsilon$. Next take $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_k| < \frac{\epsilon}{2}$ for $j, k \geq N$. Then there exists $\tilde{N} \geq N$ such that $|q_{\tilde{N}}| \geq \epsilon$. For any $j \geq N$ we then have

$$|q_j| = |q_{\tilde{N}} - (q_{\tilde{N}} - q_j)| \geq ||q_{\tilde{N}}| - |q_{\tilde{N}} - q_j|| \geq \epsilon - \tfrac{\epsilon}{2} = \tfrac{\epsilon}{2},$$

where we have used Exercise 2.2.7. The claim follows by taking δ to be the smallest of the numbers $\frac{\epsilon}{2}, |q_1|, \dots, |q_N|$.

Now let $\epsilon \in \mathbb{Q}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_k| < \delta^2\epsilon$ for $j, k \geq N$. Then

$$|q_j^{-1} - q_k^{-1}| = \left|\frac{q_k - q_j}{q_j q_k}\right| < \frac{\delta^2\epsilon}{\delta^2} = \epsilon, \qquad j, k \geq N.$$

(iv) For $\epsilon \in \mathbb{Q}_{>0}$ let $N \in \mathbb{Z}_{>0}$ have the property that $|\tilde{q}_j - q_j|, |\tilde{r}_j - r_j| < \frac{\epsilon}{2}$. Then, using the triangle inequality,

$$|(\tilde{q}_j + \tilde{r}_j) - (q_k + r_k)| \leq |\tilde{q}_j - q_k| + |\tilde{r}_k - r_k| < \epsilon, \qquad j, k \geq N.$$

(v) Let $M \in \mathbb{Q}_{>0}$ have the property that $|\tilde{q}_j|, |r_j| < M$ for all $j \in \mathbb{Z}_{>0}$. Then, for $\epsilon \in \mathbb{Q}_{>0}$, take $N \in \mathbb{Z}_{>0}$ such that $|\tilde{r}_j - r_k|, |\tilde{q}_j - q_k| < \frac{\epsilon}{2M}$ for $j, k \geq N$. We then use the triangle inequality to give

$$|(\tilde{q}_j \cdot \tilde{r}_j) - (q_k \cdot r_k)| = |\tilde{q}_j(\tilde{r}_j - r_k) - r_k(q_k - \tilde{q}_j)| < \epsilon, \qquad j, k \geq N.$$

(vi) Let $\delta \in \mathbb{Q}_{>0}$ satisfy $|q_j|, |\tilde{q}_j| \geq \delta$ for all $j \in \mathbb{Z}_{>0}$. Then, for $\epsilon \in \mathbb{Q}_{>0}$, choose $N \in \mathbb{Z}_{>0}$ such that $|\tilde{q}_j - q_j| < \delta^2\epsilon$ for $j \geq N$. Then we have

$$|\tilde{q}_j^{-1} - q_j^{-1}| = \left|\frac{q_j - \tilde{q}_j}{q_j \tilde{q}_j}\right| < \frac{\delta^2\epsilon}{\delta^2}, \qquad j \geq N,$$

so completing the proof. ■

The requirement, in parts (iii) and (vi), that the sequence $(q_j)_{j\in\mathbb{Z}_{>0}}$ have no zero elements is not really a restriction in the same way as is the requirement that the sequence not converge to zero. The reason for this is that, as we showed in the proof, if the sequence does not converge to zero, then there exists $\epsilon \in \mathbb{Q}_{>0}$ and $N \in \mathbb{Z}_{>0}$ such that $|q_j| > \epsilon$ for $j \geq N$. Thus the tail of the sequence is guaranteed to have no zero elements, and the tail of the sequence is all that matters for the equivalence class.

Now that we have shown how to add and multiply Cauchy sequences in $\mathbb{Q}$, and that this addition and multiplication depends only on equivalence classes under the notion of equivalence given in Definition 2.1.16, we can easily define addition and multiplication in $\mathbb{R}$.

2.2.2 Definition (Addition, multiplication, and division in $\mathbb{R}$) Define the operations of ***addition***, ***multiplication***, and ***division*** in $\mathbb{R}$ by

(i) $[(q_j)_{j\in\mathbb{Z}_{>0}}] + [(r_j)_{j\in\mathbb{Z}_{>0}}] = [(q_j)_{j\in\mathbb{Z}_{>0}} + (r_j)_{j\in\mathbb{Z}_{>0}}]$,

(ii) $[(q_j)_{j\in\mathbb{Z}_{>0}}] \cdot [(r_j)_{j\in\mathbb{Z}_{>0}}] = [(q_j)_{j\in\mathbb{Z}_{>0}} \cdot (r_j)_{j\in\mathbb{Z}_{>0}}]$,

(iii) $[(q_j)_{j\in\mathbb{Z}_{>0}}]/[(r_j)_{j\in\mathbb{Z}_{>0}}] = [(q_j/r_j)_{j\in\mathbb{Z}_{>0}} + (r_j)_{j\in\mathbb{Z}_{>0}}]$,

respectively, where, in the definition of division, we require that the sequence $(r_j)_{j\in\mathbb{Z}_{>0}}$ have no zero elements, and that it not converge to 0. We will sometimes omit the "·" when writing multiplication. •

Similarly to what we have done previously with $\mathbb{Z}$ and $\mathbb{Q}$, we let $-x = [(-1)_{j\in\mathbb{Z}_{>0}}]\cdot x$. For $x \in \mathbb{R}\setminus\{0\}$, we also denote by x^{-1} the real number corresponding to a Cauchy sequence $(\frac{1}{q_j})_{j\in\mathbb{Z}_{>0}}$, where $x = [(q_j)_{j\in\mathbb{Z}_{>0}}]$.

As with integers and rational numbers, we can define powers of real numbers. For $x \in \mathbb{R}\setminus\{0\}$ and $k \in \mathbb{Z}_{\geq 0}$ we define $x^k \in \mathbb{R}$ inductively by $x^0 = 1$ and $x^{k^+} = x^k \cdot x$. As usual, we call x^k the kth ***power*** of x. For $k \in \mathbb{Z}\setminus\mathbb{Z}_{\geq 0}$, we take $x^k = (x^{-k})^{-1}$. For real numbers, the notion of the power of a number can be extended. Let us show how this is done. In the statement of the result, we use the notion of positive real numbers which are not defined until Definition 2.2.8. Also, in our proof, we refer ahead to properties of $\mathbb{R}$ that are not considered until Section 2.3. However, it is convenient to state the construction here.

2.2.3 Proposition ($\mathbf{x^{1/k}}$) *For* $\mathrm{x} \in \mathbb{R}_{>0}$ *and* $\mathrm{k} \in \mathbb{Z}_{>0}$*, there exists a unique* $\mathrm{y} \in \mathbb{R}_{>0}$ *such that* $\mathrm{y^k = x}$*. We denote the number* y *by* $\mathrm{x^{1/k}}$*.*

Proof Let $S_x = \{y \in \mathbb{R} \mid y^k < x\}$. Since $x \geq 0$, $0 \in S$ so $S \neq \emptyset$. We next claim that $\max\{1, x\}$ is an upper bound for S_x. First suppose that $x < 1$. Then, for $y \in S_x$, $y^k < x < 1$, and so 1 is an upper bound for S_x. If $x \geq 1$ and $y \in S_x$, then we claim that $y \leq x$. Indeed, if $y > x$ then $y^k > x^k > x$, and so $y \notin S_x$. This shows that S_x is upper bounded by x in this case. Now we know that S_x has a least upper bound by Theorem 2.3.7. Let y denote this least upper bound.

We shall now show that $y^k = x$. Suppose that $y^k \neq x$. From Corollary 2.2.9 we have $y^k < x$ or $y^k > x$.

Suppose first that $y^k < x$. Then, for $\epsilon \in \mathbb{R}_{>0}$ we have

$$(y+\epsilon)^k = \epsilon^k + a_{k-1}y\epsilon^{k-1} + \cdots + a_1 y^{k-1}\epsilon + y^k$$

for some numbers $a_1, \dots, a_{k-1}$ (these are the binomial coefficients of Exercise 2.2.1). If $\epsilon \leq 1$ then $\epsilon^k \leq \epsilon$ for $k \in \mathbb{Z}_{>0}$. Therefore, if $\epsilon \leq 1$ we have

$$(y+\epsilon)^k \leq \epsilon(1 + a_{k-1}y + \cdots + a_1 y^{k-1}) + y^k.$$

Now, if $\epsilon < \min\{1, \frac{x-y^k}{1+a_{k-1}y+\cdots+a_a y^{k-1}}\}$, then $(y+\epsilon)^k < x$, contradicting the fact that y is an upper bound for S_x.

Now suppose that $y^k > x$. Then, for $\epsilon \in \mathbb{R}_{>0}$, we have

$$(y-\epsilon)^k = (-1)^k\epsilon^k + (-1)^{k-1}a_{k-1}y\epsilon^{k-1} + \cdots - a_1 y^{k-1}\epsilon + y^k.$$

The sum on the right involves terms that are positive and negative. This sum will be greater than the corresponding sum with the positive terms involving powers of ϵ removed. That is to say,

$$(y-\epsilon)^k > y^k - a_1 y^{k-1}\epsilon - a_3 y^{k-3}\epsilon^3 + \cdots .$$

For $\epsilon \leq 1$ we again gave $\epsilon^k \leq \epsilon$ for $k \in \mathbb{Z}_{>0}$. Therefore

$$(y-\epsilon)^k > y^k - (a_1 y^{k-1} + a_3 y^{k-3} + \cdots)\epsilon.$$

Thus, if $\epsilon < \min\{1, \frac{y^k - x}{a_1 y^{k-1} + a_3 y^{k-3} + \dots}\}$ we have $(y-\epsilon)^k > x$, contradicting the fact that y is the least upper bound for S_x.

We are forced to conclude that $y^k = x$, so giving the result. ■

If $x \in \mathbb{R}_{>0}$ and $q = \frac{j}{k} \in \mathbb{Q}$ with $j \in \mathbb{Z}$ and $k \in \mathbb{Z}_{>0}$, we define $x^q = (x^{1/k})^j$.

Let us record the basic properties of addition and multiplication, mirroring analogous results for $\mathbb{Q}$. The properties all follow easily from the similar properties for $\mathbb{Q}$, along with Proposition 2.2.1 and the definition of addition and multiplication in $\mathbb{R}$.

2.2.4 Proposition (Properties of addition and multiplication in $\mathbb{R}$) *Addition and multiplication in $\mathbb{R}$ satisfy the following rules:*

(i) $x_1 + x_2 = x_2 + x_1$, $x_1, x_2 \in \mathbb{R}$ *(**commutativity** of addition);*

(ii) $(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3)$, $x_1, x_2, x_3 \in \mathbb{R}$ *(**associativity** of addition);*

(iii) $x + 0 = x$, $t \in \mathbb{R}$ *(**additive identity**);*

(iv) $x + (-x) = 0$, $x \in \mathbb{R}$ *(**additive inverse**);*

(v) $x_1 \cdot x_2 = x_2 \cdot x_1$, $x_1, x_2 \in \mathbb{R}$ *(**commutativity** of multiplication);*

(vi) $(x_1 \cdot x_2) \cdot x_3 = x_1 \cdot (x_2 \cdot x_3)$, $x_1, x_2, x_3 \in \mathbb{R}$ *(**associativity** of multiplication);*

(vii) $x \cdot 1 = x$, $x \in \mathbb{R}$ *(**multiplicative identity**);*

(viii) $x \cdot x^{-1} = 1$, $x \in \mathbb{R} \setminus \{0\}$ *(**multiplicative inverse**);*

(ix) $y \cdot (x_1 + x_2) = y \cdot x_1 + y \cdot x_2$, $y, x_1, x_2 \in \mathbb{R}$ *(**distributivity**);*

(x) $x^{k_1} \cdot x^{k_2} = x^{k_1 + k_2}$, $x \in \mathbb{R}$, $k_1, k_2 \in \mathbb{Z}_{\geq 0}$.

Moreover, if we define $i_{\mathbb{Q}} \colon \mathbb{Q} \to \mathbb{R}$ *by* $i_{\mathbb{Q}}(q) = [(q)_{j \in \mathbb{Z}_{>0}}]$, *then addition and multiplication in $\mathbb{R}$ agrees with that in $\mathbb{Q}$:*

$$i_{\mathbb{Q}}(q_1) + i_{\mathbb{Q}}(q_2) = i_{\mathbb{Q}}(q_1 + q_2), \quad i_{\mathbb{Q}}(q_1) \cdot i_{\mathbb{Q}}(q_2) = i_{\mathbb{Q}}(q_1 \cdot q_2).$$

As we have done in the past with $\mathbb{Z} \subseteq \mathbb{Q}$, we will often regard $\mathbb{Q}$ as a subset of $\mathbb{R}$ without making explicit mention of the inclusion $i_{\mathbb{Q}}$. Note that this also allows us to think of both $\mathbb{Z}_{\geq 0}$ and $\mathbb{Z}$ as subsets of $\mathbb{R}$, since $\mathbb{Z}_{\geq 0}$ is regarded as a subset of $\mathbb{Z}$, and since $\mathbb{Z} \subseteq \mathbb{Q}$. Of course, this is nothing surprising. Indeed, perhaps the more surprising thing is that it is not actually the case that the definitions do not precisely give $\mathbb{Z}_{\geq 0} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$!

Now is probably a good time to mention that an element of $\mathbb{R}$ that is not in the image of $i_{\mathbb{Q}}$ is called ***irrational***. Also, one can show that the set $\mathbb{Q}$ of rational numbers is countable (Exercise 2.1.3), but that the set $\mathbb{R}$ of real numbers is uncountable (Exercise 2.1.4). Note that it follows that the set of irrational numbers is uncountable, since an uncountable set cannot be a union of two countable sets.

2.2.2 The total order on $\mathbb{R}$

Next we define in $\mathbb{R}$ a natural total order. To do so requires a little work. The approach we take is this. On the set $\mathsf{CS}(\mathbb{Q})$ of Cauchy sequences in $\mathbb{Q}$ we define a partial order that is *not* a total order. We then show that, for any two Cauchy sequences, in each equivalence class in $\mathsf{CS}(\mathbb{Q})$ with respect to the equivalence relation of Definition 2.1.16, there exists representatives that can be compared using the order. In this way, while the order on the set of Cauchy sequences is not a total order, there is induced a total order on the set of equivalence classes.

First we define the partial order on the set of Cauchy sequences.

2.2.5 Definition (Partial order on CS($\mathbb{Q}$)) The partial order $\preceq$ on $\mathsf{CS}(\mathbb{Q})$ is defined by

$$(q_j)_{j\in\mathbb{Z}_{>0}} \preceq (r_j)_{j\in\mathbb{Z}_{>0}} \quad\iff\quad q_j \leq r_j,\ j\in\mathbb{Z}_{>0}.$$

•

This partial order is clearly not a total order. For example, the Cauchy sequences $(\frac{1}{j})_{j\in\mathbb{Z}_{>0}}$ and $(\frac{(-1)^j}{j})_{j\in\mathbb{Z}_{>0}}$ are not comparable with respect to this order. However, what is true is that equivalence classes of Cauchy sequences *are* comparable. We refer the reader to Definition 2.1.16 for the definition of the equivalence relation we denote by $\sim$ in the following result.

2.2.6 Proposition *Let* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \mathsf{CS}(\mathbb{Q})$ *and suppose that* $(\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \nsim (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$. *The following two statements hold:*

(i) There exists $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \mathsf{CS}(\mathbb{Q})$ *such that*

(a) $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$, *and*

(b) either $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \prec (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *or* $(\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \prec (\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$.

(ii) There does not exist $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\bar{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\bar{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \mathsf{CS}(\mathbb{Q})$ *such that*

(a) $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\bar{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\mathrm{q_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\bar{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \sim (\mathrm{r_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$, *and*

(b) one of the following two statements holds:

I. $(\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \prec (\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\bar{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \prec (\bar{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$;

II. $(\tilde{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \prec (\tilde{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\bar{\mathrm{q}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \prec (\bar{\mathrm{r}}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$.

Proof (i) We begin with a useful lemma.

1 Lemma *With the given hypotheses, there exists* $\delta\in\mathbb{Q}_{>0}$ *and* $\mathrm{N}\in\mathbb{Z}_{>0}$ *such that* $|\mathrm{q_j}-\mathrm{r_j}|\geq\delta$ *for all* $\mathrm{j}\geq\mathrm{N}$.

Proof Since $(q_j - r_j)_{j\in\mathbb{Z}_{>0}}$ does not converge to zero, choose $\epsilon\in\mathbb{Q}_{>0}$ such that, for all $N\in\mathbb{Z}_{>0}$, there exists $j\geq N$ such that $|q_j - r_j|\geq\epsilon$. Now take $N\in\mathbb{Z}_{>0}$ such that $|q_j-q_k|, |r_k-r_k|\leq\frac{\epsilon}{4}$ for $j,k\geq N$. Then, by our assumption about ϵ, there exists $\tilde{N}\geq N$ such that $|q_{\tilde{N}} - r_{\tilde{N}}|\geq\epsilon$. Then, for any $j\geq N$, we have

$$\begin{aligned}|q_j - r_j| &= |(q_{\tilde{N}} - r_{\tilde{N}}) - (q_{\tilde{N}} - r_{\tilde{N}}) - (q_j - r_j)|\\ &\geq ||q_{\tilde{N}} - r_{\tilde{N}}| - |(q_{\tilde{N}} - r_{\tilde{N}}) - (q_j - r_j)|| \geq \epsilon - \tfrac{\epsilon}{2}.\end{aligned}$$

The lemma follows by taking $\delta = \frac{\epsilon}{2}$. ▼

Now take N and δ as in the lemma. Then take $\tilde{N} \in \mathbb{Z}_{>0}$ such that $|q_j - q_k|, |r_j - r_k| < \frac{\delta}{2}$ for $j, k \geq \tilde{N}$. Then, using the triangle inequality,

$$|(q_j - r_j) - (q_k - r_k)| \leq \delta, \qquad j, k \geq \tilde{N}.$$

Now take K to be the larger of N and $\tilde{N}$. We then have either $q_K - r_K \geq \delta$ or $r_K - q_K \geq \delta$. First suppose that $q_K - r_K \geq \delta$ and let $j \geq K$. Either $q_j - r_j \geq \delta$ or $r_j - q_j \geq \delta$. If the latter, then

$$q_j - r_j \leq -\delta \quad \Longrightarrow \quad (q_j - r_k) - (q_K - r_K) \leq 2\delta,$$

contradicting the definition of K. Therefore, we must have $q_j - r_j \geq \delta$ for all $j \geq K$. A similar argument when $r_K - q_K \geq \delta$ shows that $r_j - q_j \geq \delta$ for all $j \geq K$. For $j \in \mathbb{Z}_{>0}$ we then define

$$\tilde{q}_j = \begin{cases} q_K, & j < K, \\ q_j, & j \geq K, \end{cases} \qquad \tilde{r}_j = \begin{cases} r_K, & j < K, \\ r_j, & j \geq K, \end{cases},$$

and we note that the sequences $(\tilde{q}_j)_{j\in\mathbb{Z}_{>0}}$ and $(\tilde{r}_j)_{j\in\mathbb{Z}_{>0}}$ satisfy the required conditions.

(ii) Suppose that

1. $(q_j)_{j\in\mathbb{Z}_{>0}} \nsim (r_j)_{j\in\mathbb{Z}_{>0}}$,
2. $(\tilde{q}_j)_{j\in\mathbb{Z}_{>0}} \sim (\bar{q}_j)_{j\in\mathbb{Z}_{>0}} \sim (q_j)_{j\in\mathbb{Z}_{>0}}$,
3. $(\tilde{r}_j)_{j\in\mathbb{Z}_{>0}} \sim (\bar{r}_j)_{j\in\mathbb{Z}_{>0}} \sim (r_j)_{j\in\mathbb{Z}_{>0}}$, and
4. $(\tilde{q}_j)_{j\in\mathbb{Z}_{>0}} \prec (\tilde{r}_j)_{j\in\mathbb{Z}_{>0}}$.

From the previous part of the proof we know that there exists $\delta \in \mathbb{Q}_{>0}$ and $N \in \mathbb{Z}_{>0}$ such that $\tilde{q}_j - \tilde{r}_j \geq \delta$ for $j \geq N$. Then take $\tilde{N} \in \mathbb{Z}_{>0}$ such that $|\tilde{q}_j - \bar{q}_j|, |\tilde{r}_j - \bar{r}_j| < \frac{\delta}{4}$ for $j \geq \tilde{N}$. This implies that for $j \geq \tilde{N}$ we have

$$|(\tilde{q}_j - \tilde{r}_j) - (\bar{q}_j - \bar{r}_j)| < \tfrac{\delta}{2}.$$

Therefore,

$$(\bar{q}_j - \bar{r}_j) > (\tilde{q}_j - \tilde{r}_j) - \tfrac{\delta}{2}, \qquad j \geq \tilde{N}.$$

If additionally $j \geq N$, then we have

$$(\bar{q}_j - \bar{r}_j) > \delta - \tfrac{\delta}{2} = \tfrac{\delta}{2}.$$

This shows the impossibility of $(\bar{r}_j)_{j\in\mathbb{Z}_{>0}} \prec (\bar{q}_j)_{j\in\mathbb{Z}_{>0}}$. A similar argument shows that $(\tilde{r}_j)_{j\in\mathbb{Z}_{>0}} \prec (\tilde{q}_j)_{j\in\mathbb{Z}_{>0}}$ bars the possibility that $(\bar{q}_j)_{j\in\mathbb{Z}_{>0}} \prec (\bar{r}_j)_{j\in\mathbb{Z}_{>0}}$. ■

Using the preceding result, the following definition then makes sense.

2.2.7 Definition (Order on $\mathbb{R}$) The total order on $\mathbb{R}$ is defined by $x \leq y$ if and only if there exists $(q_j)_{j\in\mathbb{Z}_{>0}}, (r_j)_{j\in\mathbb{Z}_{>0}} \in \mathrm{CS}(\mathbb{Q})$ such that

(i) $x = [(q_j)_{j\in\mathbb{Z}_{>0}}]$ and $y = [(r_j)_{j\in\mathbb{Z}_{>0}}]$ and

(ii) $(q_j)_{j\in\mathbb{Z}_{>0}} \preceq (r_j)_{j\in\mathbb{Z}_{>0}}$. •

Note that we have used the symbol "$\leq$" for the total order on $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. This is justified since, if we think of $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$, then the various total orders agree (Exercises 2.1.2 and 2.2.5).

We have the usual language and notation we associate with various kinds of numbers.

2.2.8 Definition (Positive and negative real numbers) A real number x is:

(i) ***positive*** if $0 < x$;

(ii) ***negative*** if $x < 0$;

(iii) ***nonnegative*** if $0 \leq x$;

(iv) ***nonpositive*** if $x \leq 0$.

The set of positive real numbers is denoted by $\mathbb{R}_{>0}$, the set of nonnegative real numbers is denoted by $\mathbb{R}_{\geq 0}$, the set of negative real numbers is denoted by $\mathbb{R}_{<0}$, and the set of nonpositive real numbers is denoted by $\mathbb{R}_{\leq 0}$. •

Now is a convenient moment to introduce some simple notation and concepts that are associated with the natural total order on $\mathbb{R}$. The ***signum function*** is the map $\operatorname{sign}\colon \mathbb{R} \to \{-1, 0, 1\}$ defined by

$$\operatorname{sign}(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0. \end{cases}$$

For $x \in \mathbb{R}$, $\lceil x \rceil$ is the ***ceiling*** of x which is the smallest integer not less than x. Similarly, $\lfloor x \rfloor$ is the ***floor*** of x which is the largest integer less than or equal to x. In Figure 2.1 we show the ceiling and floor functions.

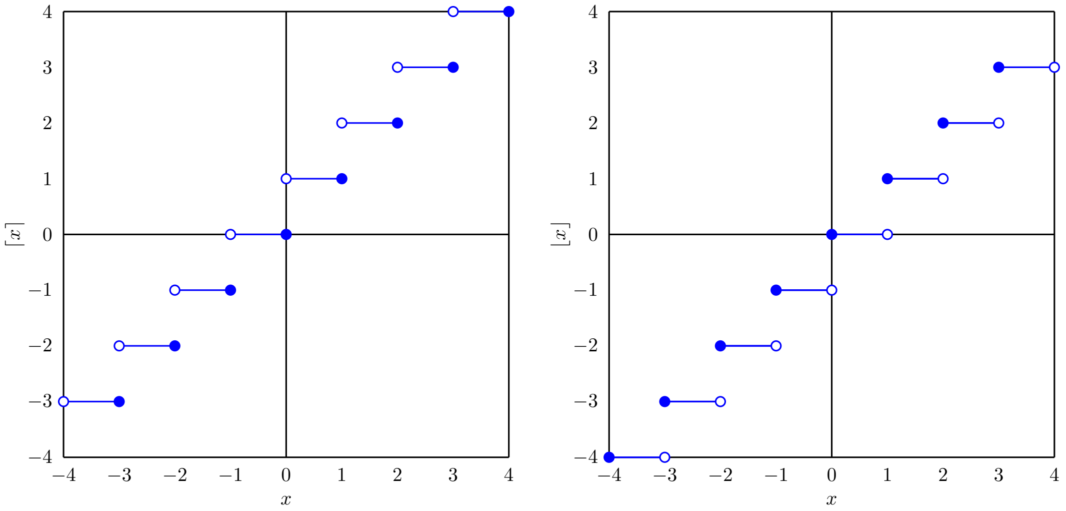


Figure 2.1 The ceiling function (left) and floor function (right)

A consequence of our definition of order is the following extension of the Trichotomy Law to $\mathbb{R}$.

2.2.9 Corollary (Trichotomy Law for $\mathbb{R}$) *For $x, y \in \mathbb{R}$, exactly one of the following possibilities holds:*

(i) $x < y$;

(ii) $y < x$;

(iii) $x = y$.

As with integers and rational numbers, addition and multiplication of real numbers satisfy the expected properties with respect to the total order.

2.2.10 Proposition (Relation between addition and multiplication and <) *For* $x, y, z \in \mathbb{R}$, *the following statements hold:*

(i) *if* $x < y$ *then* $x + z < y + z$;

(ii) *if* $x < y$ *and if* $z > 0$ *then* $z \cdot x < z \cdot y$;

(iii) *if* $x < y$ *and if* $z < 0$ *then* $z \cdot y < z \cdot x$;

(iv) *if* $0 < x, y$ *then* $0 < x \cdot y$;

(v) *if* $x < y$ *and if either*

(a) $0 < x, y$ *or*

(b) $x, y < 0$,

then $y^{-1} < x^{-1}$.

Proof These statements all follow from the similar statements for $\mathbb{Q}$, along with Proposition 2.2.6. We leave the straightforward verifications to the reader as Exercise 2.2.4. ■

2.2.3 The absolute value function on $\mathbb{R}$

In this section we generalise the absolute value function on $\mathbb{Q}$. As we shall see in subsequent sections, this absolute value function is essential for providing much of the useful structure of the set of real numbers.

The definition of the absolute value is given as usual.

2.2.11 Definition (Real absolute value function) The ***absolute value function*** on $\mathbb{R}$ is the map from $\mathbb{R}$ to $\mathbb{R}_{\geq 0}$, denoted by $x \mapsto |x|$, defined by

$$|x| = \begin{cases} x, & 0 < x, \\ 0, & x = 0, \\ -x, & x < 0. \end{cases} \qquad \bullet$$

Note that we have used the symbol "$|\cdot|$" for the absolute values on $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. This is justified since, if we think of $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$, then the various absolute value functions agree (Exercises 2.1.2 and 2.2.5).

The real absolute value function has the expected properties. The proof of the following result is straightforward, and so omitted.

2.2.12 Proposition (Properties of absolute value on $\mathbb{R}$) *The following statements hold:*

(i) $|x| \geq 0$ *for all* $x \in \mathbb{R}$;

(ii) $|x| = 0$ *if and only if* $x = 0$;

(iii) $|x \cdot y| = |x| \cdot |y|$ *for all* $x, y \in \mathbb{R}$;

(iv) $|x + y| \leq |x| + |y|$ *for all* $x, y \in \mathbb{R}$ *(**triangle inequality**)*;

(v) $|x^{-1}| = |x|^{-1}$ *for all* $x \in \mathbb{R} \setminus \{0\}$.

2.2.4 Properties of $\mathbb{Q}$ as a subset of $\mathbb{R}$

In this section we give some seemingly obvious, and indeed not difficult to prove, properties of the rational numbers as a subset of the real numbers.

The first property bears the name of Archimedes,[2] but Archimedes actually attributes this to Eudoxus.[3] In any case, it is an Ancient Greek property.

2.2.13 Proposition (Archimedean property of $\mathbb{R}$) *Let $\epsilon \in \mathbb{R}_{>0}$. Then, for any $x \in \mathbb{R}$ there exists $k \in \mathbb{Z}_{>0}$ such that $k \cdot \epsilon > x$.*

Proof Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ and $(e_j)_{j\in\mathbb{Z}_{>0}}$ be Cauchy sequences in $\mathbb{Q}$ such that $x = [(q_j)_{j\in\mathbb{Z}_{>0}}]$ and $\epsilon = [(e_j)_{j\in\mathbb{Z}_{>0}}]$. By Proposition 2.1.14 there exists $M \in \mathbb{R}_{>0}$ such that $|q_j| < M$ for all $j \in \mathbb{Z}_{>0}$, and by Proposition 2.2.6 we may suppose that $e_j > \delta$ for $j \in \mathbb{Z}_{>0}$, for some $\delta \in \mathbb{Q}_{>0}$. Let $k \in \mathbb{Z}_{>0}$ satisfy $k > \frac{M+1}{\delta}$ (why is this possible?). Then we have

$$k \cdot e_j > \frac{M+1}{\delta} \cdot \delta = M + 1 \geq q_j + 1, \qquad j \in \mathbb{Z}_{>0}.$$

Now consider the sequence $(k \cdot e_j - q_j)_{j\in\mathbb{Z}_{>0}}$. This is a Cauchy sequence by Proposition 2.2.1 since it is a sum of products of Cauchy sequences. Moreover, our computations show that each term in the sequence is larger than 1. Also, this Cauchy sequence has the property that $[(k \cdot e_j - q_j)_{j\in\mathbb{Z}_{>0}}] = k \cdot \epsilon - x$. This shows that $k \cdot \epsilon - x \in \mathbb{R}_{>0}$, so giving the result. ■

The Archimedean property roughly says that there are no real numbers which are greater all rational numbers. The next result says that there are no real numbers that are smaller than all rational numbers.

2.2.14 Proposition (There is no smallest positive real number) *If $\epsilon \in \mathbb{R}_{>0}$ then there exists $q \in \mathbb{Q}_{>0}$ such that $q < \epsilon$.*

Proof Since $\epsilon^{-1} \in \mathbb{R}_{>0}$ let $k \in \mathbb{Z}_{>0}$ satisfy $k \cdot 1 > \epsilon^{-1}$ by Proposition 2.2.13. Then taking $q = k^{-1} \in \mathbb{Q}_{>0}$ gives $q < \epsilon$. ■

Using the preceding two results, it is then easy to see that arbitrarily near any real number lies a rational number.

2.2.15 Proposition (Real numbers are well approximated by rational numbers I) *If $x \in \mathbb{R}$ and if $\epsilon \in \mathbb{R}_{>0}$, then there exists $q \in \mathbb{Q}$ such that $|x - q| < \epsilon$.*

Proof If $x = 0$ then the result follows by taking $q = 0$. Let us next suppose that $x > 0$. If $x < \epsilon$ then the result follows by taking $q = 0$, so we assume that $x \geq \epsilon$. Let $\delta \in \mathbb{Q}_{>0}$ satisfy $\delta < \epsilon$ by Proposition 2.2.14. Then use Proposition 2.2.13 to choose $k \in \mathbb{Z}_{>0}$ to satisfy $k \cdot \delta > x$. Moreover, since $x > 0$, we will assume that k is the smallest such

[2] Archimedes of Syracuse (287 BC–212 BC) was a Greek mathematician and physicist (although in that era such classifications of scientific aptitude were less rigid than they are today). Much of his mathematical work was in the area of geometry, but many of Archimedes' best known achievements were in physics (e.g., the Archimedean Principle in fluid mechanics). The story goes that when the Romans captured Syracuse in 212 BC, Archimedes was discovered working on some mathematical problem, and struck down in the act by a Roman soldier.

[3] Eudoxus of Cnidus (408 BC–355 BC) was a Greek mathematician and astronomer. His mathematical work was concerned with geometry and numbers.

number. Since $x \geq \epsilon$, $k \geq 2$. Thus $(k-1)\cdot\delta \leq x$ since k is the smallest natural number for which $k \cdot \delta > x$. Now we compute

$$0 \leq x - (k-1)\cdot\delta < k\cdot\delta - (k-1)\cdot\delta = \delta < \epsilon.$$

It is now easy to check that the result holds by taking $q = (k-1)\cdot\delta$. The situation when $x < 0$ is easily shown to follow from the situation when $x > 0$. ■

The following stronger result is also useful, and can be proved along the same lines as Proposition 2.2.15, using the Archimedean property of $\mathbb{R}$. The reader is asked to do this as Exercise 2.2.3.

2.2.16 Corollary (Real numbers are well approximated by rational numbers II) *If* $x, y \in \mathbb{R}$ *with* $x < y$*, then there exists* $q \in \mathbb{Q}$ *such that* $x < q < y$*.*

One can also show that irrational numbers have the same property.

2.2.17 Proposition (Real numbers are well approximated by irrational numbers) *If* $x \in \mathbb{R}$ *and if* $\epsilon \in \mathbb{R}_{>0}$*, then there exists* $y \in \mathbb{R} \setminus \mathbb{Q}$ *such that* $|x - y| < \epsilon$*.*

Proof By Corollary 2.2.16 choose $q_1, q_2 \in \mathbb{Q}$ such that $x - \epsilon < q_1 < q_2 < x + \epsilon$. Then the number

$$y = q_1 + \frac{q_2 - q_1}{\sqrt{2}}$$

is irrational and satisfies $q_1 < y < q_2$. Therefore, $x - \epsilon < y < x + \epsilon$, or $|x - y| < \epsilon$. ■

It is also possible to state a result regarding the approximation of a collection of real numbers by rational numbers of a certain form. The following result gives one such result.

2.2.18 Theorem (Dirichlet Simultaneous Approximation Theorem) *If* $x_1, \dots, x_k \in \mathbb{R}$ *and if* $N \in \mathbb{Z}_{>0}$*, then there exists* $m \in \{1, \dots, N^k\}$ *and* $m_1, \dots, m_k \in \mathbb{Z}$ *such that*

$$\max\{|mx_1 - m_1|, \dots, |mx_k - m_k|\} < \frac{1}{N}.$$

Proof Let

$$C = [0,1)^k \subseteq \mathbb{R}^k$$

be the "cube" in $\mathbb{R}^k$. For $j \in \{1, \dots, N\}$ denote $I_j = [\frac{j-1}{N}, \frac{j}{N})$ and note that the sets

$$\{I_{j_1} \times \cdots \times I_{j_k} \subseteq C \mid j_1, \dots, j_k \in \{1, \dots, N\}\}$$

form a partition of the cube C into N^k "subcubes." Now consider the $N^k + 1$ points

$$\{(lx_1, \dots, lx_k) \mid l \in \{0, 1, \dots, N^k\}\}$$

in $\mathbb{R}^k$. If $\lfloor x \rfloor$ denotes the floor of $x \in \mathbb{R}$ (i.e., the largest integer less than or equal to x), then

$$\{(lx_1 - \lfloor lx_1 \rfloor, \dots, lx_k - \lfloor lx_k \rfloor) \mid l \in \{0, 1, \dots, N^k\}\}$$

is a collection of $N^k + 1$ numbers in C. Since C is partitioned into the N^k cubes, it must be that at least two of these $N^k + 1$ points lie in the same cube. Let these points correspond to $l_1, l_2 \in \{0, 1, \dots, n^k\}$ with $l_2 > l_1$. Then, letting $m = l_2 - l_2$ and $m_j = \lfloor l_2 x_j \rfloor - \lfloor l_1 x_j \rfloor$, $j \in \{1, \dots, k\}$, we have

$$|mx_j - m_j| = |l_2 - \lfloor l_2 x_j \rfloor - (l_1 x_j - \lfloor l_1 x_j \rfloor)| < \frac{1}{N}$$

for every $j \in \{1, \dots, k\}$, which is the result since $m \in \{1, \dots, N^k\}$. ■

2.2.19 Remark (Dirichlet's "pigeonhole principle") The proof of the preceding theorem is a clever application of the so-called "pigeonhole principle," whose use seems to have been pioneered by Dirichlet. The idea behind this principle is simple. One uses the problem data to define elements $x_1, \dots, x_m$ of some set S. One then constructs a partition $(S_1, \dots, S_k)$ of S with the property that, if any $x_{j_1}, x_{j_2} \in S_l$ for some $l \in \{1, \dots, k\}$ and some $j_1, j_2 \in \{1, \dots, m\}$, then the desired result holds. If $k > m$ this is automatically satisfied. •

Note that the previous result gives an arbitrarily accurate simultaneous approximation of the numbers $x_1, \dots, x_j$ by rational numbers with the same denominator since we have

$$\left| x_j - \frac{m_j}{m} \right| < \frac{1}{mN^k} \le \frac{1}{N^{k+1}}.$$

By choosing N large, our simultaneous approximations can be made as good as desired.

Let us now ask a somewhat different sort of question. Given a fixed set $a_1, \dots, a_k \in \mathbb{R}$, what are the conditions on these numbers such that, given *any* set $x_1, \dots, x_k \in \mathbb{R}$, we can find another number $b \in \mathbb{R}$ such that the approximations $|ba_j - x_j|$, $j \in \{1, \dots, k\}$, are arbitrarily close to integer multiples of a certain number. The exact reason why this is interesting is not immediately clear, but becomes clear in Theorem **??** when we talk about the geometry of the unit circle in the complex plane. In any event, the following result addresses this approximation question, making reference to the notion of linear independence which we discuss in Section 4.3.3. In the statement of the theorem, we think of $\mathbb{R}$ as being a $\mathbb{Q}$-vector space.

2.2.20 Theorem (Kronecker Approximation Theorem) *For* $a_1, \dots, a_k \in \mathbb{R}$ *and* $\Delta \in \mathbb{R}_{>0}$ *the following statements hold:*

(i) if $\{a_1, \dots, a_k\}$ *are linearly over* $\mathbb{Q}$ *then, for any* $x_1, \dots, x_k \in \mathbb{R}$, *for any* $\epsilon \in \mathbb{R}_{>0}$ *and for any* $N \in \mathbb{Z}_{>0}$, *there exists* $b \in \mathbb{R}$ *with* $b > N$ *and integers* $m_1, \dots, m_k$ *such that*

$$\max\{|ba_1 - x_1 - m_1\Delta|, \dots, |ba_k - x_k - m_k\Delta|\} < \epsilon;$$

(ii) if $\{\Delta, a_1, \dots, a_k\}$ *are linearly over* $\mathbb{Q}$ *then, for any* $x_1, \dots, x_k \in \mathbb{R}$, *for any* $\epsilon \in \mathbb{R}_{>0}$, *and for any* $N \in \mathbb{Z}_{>0}$, *there exists* $b \in \mathbb{Z}$ *with* $b > N$ *and integers* $m_1, \dots, m_k$ *such that*

$$\max\{|ba_1 - x_1 - m_1\Delta|, \dots, |ba_k - x_k - m_k\Delta|\} < \epsilon.$$

Proof Let us first suppose that $\Delta = 1$.

We prove the two assertions together, using induction on k.

First we prove (i) for $k = 1$. Thus suppose that $\{a_1\} \neq \{0\}$. Let $x_1 \in \mathbb{R}$, let $\epsilon \in \mathbb{R}_{>0}$, and let $N \in \mathbb{Z}_{>0}$. If m_1 is an integer greater than N and if $b = a_1^{-1}(x_1 + m_1)$, then we have $ba_1 - x_1 - m_1 = 0$, giving the result in this case.

Next we prove that if (i) holds for $k = r$ then (ii) also holds for $k = r$. Thus suppose that $\{1, a_1, \dots, a_r\}$ are linearly independent over $\mathbb{Q}$. Let $x_1, \dots, x_r \in \mathbb{R}$, let $\epsilon \in \mathbb{R}_{>0}$, and let $N \in \mathbb{Z}_{>0}$. By the Dirichlet Simultaneous Approximation Theorem, let $m, m'_1, \dots, m'_r \in \mathbb{Z}$ with $m \in \mathbb{Z}_{>0}$ be such that

$$|ma_j - m'_j| < \frac{\epsilon}{2}, \qquad j \in \{1, \dots, r\}.$$

We claim that $\{ma_1 - m'_1, \dots, ma_r - m'_r\}$ are linearly independent over $\mathbb{Q}$. Indeed, suppose that

$$q_1(ma_1 - m'_1) + \dots + q_r(ma_r - m'_r) = 0$$

for some $q_1, \dots, q_r \in \mathbb{Q}$. Then we have

$$(mq_1)a_1 + \dots + (mq_r)a_r) - (m'_1q_1 + \dots + m'_rq_r)1 = 0.$$

By linear independence of $\{1, a_1, \dots, a_r\}$ over $\mathbb{Q}$ it follows that $mq_j = 0$, $j \in \{1, \dots, r\}$, and so $q_j = 0$, $j \in \{1, \dots, r\}$, giving the desired linear independence. Since $\{ma_1 - m'_1, \dots, ma_r - m'_r\}$ are linearly independent over $\mathbb{Q}$, we may use our assumption that (i) holds for $k = r$ to give the existence of $b' \in \mathbb{R}$ with $b' > N + 1$ and integers $m''_1, \dots, m''_r$ such that

$$|b'(ma_j - m'_j) - x_j - m''_j| < \frac{\epsilon}{2}, \qquad j \in \{1, \dots, r\}.$$

Now let $b = \lfloor b' \rfloor m > N$ and $m_j = m''_j + \lfloor b' \rfloor m'_j$, $j \in \{1, \dots, k\}$. Using the triangle inequality we have

$$\begin{aligned}
|ba_j - x_j - m_j| &= |\lfloor b'm \rfloor a_j - x_j - (m''_j + \lfloor b' \rfloor m'_j)| \\
&= |\lfloor b' \rfloor (ma_j - m'_j) - x_j - m''_j| \\
&= |(\lfloor b' \rfloor - b')(ma_j - m'_j) + b'(ma_j - m'_j) - x_j - m''_j| \\
&\leq |(\lfloor b' \rfloor - b')(ma_j - m'_j)| + |b'(ma_j - m'_j) - x_j - m''_j| < \epsilon,
\end{aligned}$$

as desired.

Now we prove that (ii) with $k = r$ implies (i) with $k = r + 1$. Thus let $a_1, \dots, a_{r+1}$ be linearly independent over $\mathbb{Q}$. Let $x_1, \dots, x_{r+1} \in \mathbb{R}$, let $\epsilon \in \mathbb{R}_{>0}$, and let $N \in \mathbb{Z}_{>0}$. Note that linear independence implies that $a_{r+1} \neq 0$ (see Proposition 4.3.19(ii)). We claim that $\{1, \frac{a_1}{a_{r+1}}, \dots, \frac{a_r}{a_{r+1}}\}$ are linearly independent over $\mathbb{Q}$. Since (ii) holds for $k = r$ there exists $b' \in \mathbb{Z}$ with $b' > N$ and integers $m'_1, \dots, m'_r$ such that

$$\Big| b' \frac{a_j}{a_{r+1}} - \Big(x_j - x_{r+1} \frac{a_j}{a_{r+1}}\Big) - m'_j \Big| < \epsilon, \qquad j \in \{1, \dots, r\}.$$

Rewriting this as

$$\Big| \Big(\frac{b' + x_{r+1}}{a_{r+1}}\Big) a_j - x_j - m'_j \Big| < \epsilon, \qquad j \in \{1, \dots, r\},$$

and noting that

$$\Big(\frac{b' + x_{r+1}}{a_{r+1}}\Big)a_{r+1} - x_{r+1} - b' = 0,$$

which gives (i) by taking

$$b = \frac{b' + x_{r+1}}{a_{r+1}},\ m_1 = m'_1,\ \dots,\ m_r = m'_r,\ m_{r+1} = b'.$$

The above induction arguments give the theorem with $\Delta = 1$. Now let us relax the assumption that $\Delta = 1$. Thus let $\Delta \in \mathbb{R}_{>0}$. Let us define $a'_j = \Delta^{-1}a_j$, $j \in \{1, \dots, k\}$. We claim that $\{a'_1, \dots, a'_k\}$ is linearly independent over $\mathbb{Q}$ if $\{a_1, \dots, a_k\}$ is linearly independent over $\mathbb{Q}$. Indeed, suppose that

$$q_1 a'_1 + \dots + q_k a'_k = 0$$

for some $q_1, \dots, q_k \in \mathbb{Q}$. Multiplying by Δ and using the linear independence of $\{a_1, \dots, a_k\}$ immediately gives $q_j = 0$, $j \in \{1, \dots, k\}$. We also claim that $\{1, a'_1, \dots, a'_k\}$ is linearly independent over $\mathbb{Q}$ if $\{\Delta, a_1, \dots, a_k\}$ is linearly independent over $\mathbb{Q}$. Indeed, suppose that

$$q_0\,1 + q_1 a'_1 + \dots + q_k a'_k = 0$$

for some $q_0, q_1, \dots, q_k \in \mathbb{Q}$. Multiplying by Δ and using the linear independence of $\{\Delta, a_1, \dots, a_k\}$ immediately gives $q_j = 0$, $j \in \{1, \dots, k\}$. Let $x_1, \dots, x_k \in \mathbb{R}$, $\epsilon \in \mathbb{R}_{>0}$, and $N \in \mathbb{Z}$. Define $x'_j = \Delta^{-1}x_j$, $j \in \{1, \dots, k\}$. Since the theorem holds for $\Delta = 1$, there exists $b > N$ (with $b \in \mathbb{R}$ for part (i) and $b \in \mathbb{Z}$ for part (ii)) such that

$$|ba'_j - x'_j - m_1| < \frac{\epsilon}{\Delta}, \qquad j \in \{1, \dots, k\}.$$

Multiplying the inequality by Δ gives the result. ■

2.2.5 The extended real line

It is sometimes convenient to be able to talk about the concept of "infinity" in a somewhat precise way. We do so by using the following idea.

2.2.21 Definition (Extended real line) The ***extended real line*** is the set $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$, and we denote this set by $\overline{\mathbb{R}}$. •

Note that in this definition the symbols "$-\infty$" and "∞" are to simply be thought of as labels given to the elements of the singletons $\{-\infty\}$ and $\{\infty\}$. That they somehow correspond to our ideas of what "infinity" means is a consequence of placing some additional structure on $\overline{\mathbb{R}}$, as we now describe.

First we define "arithmetic" in $\overline{\mathbb{R}}$. We can also define some rules for arithmetic in $\overline{\mathbb{R}}$.

2.2.22 Definition (Addition and multiplication in $\overline{\mathbb{R}}$) For $x, y \in \overline{\mathbb{R}}$, define

$$x + y = \begin{cases} x + y, & x, y \in \mathbb{R}, \\ \infty, & x \in \mathbb{R},\ y = \infty,\ \text{or } x = \infty,\ y \in \mathbb{R}, \\ \infty, & x = y = \infty, \\ -\infty, & x = -\infty,\ y \in \mathbb{R} \text{ or } x \in \mathbb{R},\ y = -\infty, \\ -\infty, & x = y = -\infty. \end{cases}$$

The operations $\infty + (-\infty)$ and $(-\infty) + \infty$ are undefined. Also define

$$x \cdot y = \begin{cases} x \cdot y, & x, y \in \mathbb{R}, \\ \infty, & x \in \mathbb{R}_{>0},\ y = \infty,\ \text{or } x = \infty,\ y \in \mathbb{R}_{>0}, \\ \infty, & x \in \mathbb{R}_{<0},\ y = -\infty,\ \text{or } x = -\infty,\ y \in \mathbb{R}_{<0}, \\ \infty, & x = y = \infty,\ \text{or } x = y = -\infty, \\ -\infty, & x \in \mathbb{R}_{>0},\ y = -\infty,\ \text{or } x = -\infty,\ y \in \mathbb{R}_{>0}, \\ -\infty, & x \in \mathbb{R}_{<0},\ y = \infty,\ \text{or } x = \infty,\ y \in \mathbb{R}_{<0}, \\ -\infty, & x = \infty,\ y = -\infty \text{ or } x = -\infty,\ y = \infty, \\ 0, & x = 0, y \in \{-\infty, \infty\} \text{ or } x \in \{-\infty, \infty\},\ y = 0. \end{cases} \qquad \bullet$$

2.2.23 Remarks (Algebra in $\overline{\mathbb{R}}$)

1. The above definitions of addition and multiplication on $\overline{\mathbb{R}}$ *do not* make this a field. Thus, in some sense, the operations are simply notation, since they do not have the usual properties we associate with addition and multiplication.
2. Note we *do* allow multiplication between 0 and $-\infty$ and ∞. This convention is not universally agreed upon, but it will be useful for us to do adopt this convention in Chapter 5. •

2.2.24 Definition (Order on $\overline{\mathbb{R}}$) For $x, y \in \overline{\mathbb{R}}$, write

$$x \le y \quad \iff \quad \begin{cases} x = y, & \text{or} \\ x, y \in \mathbb{R},\ x \le y, & \text{or} \\ x \in \mathbb{R},\ y = \infty, & \text{or} \\ x = -\infty,\ y \in \mathbb{R}, & \text{or} \\ x = -\infty,\ y = \infty. \end{cases} \qquad \bullet$$

This is readily verified to be a total order on $\overline{\mathbb{R}}$, with $-\infty$ being the least element and ∞ being the greatest element of $\overline{\mathbb{R}}$. As with $\mathbb{R}$, we have the notation

$$\overline{\mathbb{R}}_{>0} = \{x \in \overline{\mathbb{R}} \mid x > 0\}, \quad \overline{\mathbb{R}}_{\ge 0} = \{x \in \overline{\mathbb{R}} \mid x \ge 0\}.$$

Finally, we can extend the absolute value on $\mathbb{R}$ to $\overline{\mathbb{R}}$.

2.2.25 Definition (Extended real absolute value function) The ***extended real absolute function*** is the map from $\overline{\mathbb{R}}$ to $\overline{\mathbb{R}}_{\geq 0}$, denoted by $x \mapsto |x|$, and defined by

$$|x| = \begin{cases} |x|, & x \in \mathbb{R}, \\ \infty, & x = \infty, \\ \infty, & x = -\infty. \end{cases} \qquad \bullet$$

2.2.6 sup and inf

We recall from Definition **??** the notation $\sup S$ and $\inf S$ for the least upper bound and greatest lower bound, respectively, associated to a partial order. This construction applies, in particular to the partially ordered set $(\overline{\mathbb{R}}, \leq)$. Note that if $A \subseteq \mathbb{R}$ then we might possibly have $\sup(A) = \infty$ and/or $\inf(A) = -\infty$. In brief section we give a few properties of sup and inf.

The following property of sup and inf is often useful.

2.2.26 Lemma (Property of sup and inf) *Let* $A \subseteq \mathbb{R}$ *be such that* $\inf(A), \sup(A) \in \mathbb{R}$ *and let* $\epsilon \in \mathbb{R}_{>0}$*. Then there exists* $x_+, x_- \in A$ *such that*

$$x_+ + \epsilon > \sup(A), \quad x_- - \epsilon < \inf(A).$$

Proof We prove the assertion for sup, as the assertion for inf follows along similar lines, of course. Suppose that there is no $x_+ \in A$ such that $x_+ + \epsilon > \sup(A)$. Then $x \leq \sup(A) - \epsilon$ for every $x \in A$, and so $\sup(A) - \epsilon$ is an upper bound for A. But this contradicts $\sup(A)$ being the least upper bound. ■

Let us record and prove the properties of interest for sup.

2.2.27 Proposition (Properties of sup) *For subsets* $A, B \subseteq \mathbb{R}$ *and for* $a \in \mathbb{R}_{>0}$*, the following statements hold:*

(i) *if* $A + B = \{x + y \mid x \in A, \ y \in B\}$*, then* $\sup(A + B) = \sup(A) + \sup(B)$*;*

(ii) *if* $-A = \{-x \mid x \in A\}$*, then* $\sup(-A) = -\inf(A)$*;*

(iii) *if* $aA = \{ax \mid x \in A\}$*, then* $\sup(aA) = a\sup(A)$*;*

(iv) *if* $I \subseteq \mathbb{R}$ *is an interval, if* $A \subseteq \mathbb{R}$*, if* $f\colon I \to \mathbb{R}$ *is strictly monotonically (see Definition 3.1.27), and if* $f(A) = \{f(x) \mid x \in A\}$*, then* $\sup(f(A)) = f(\sup(A))$*.*

Proof (i) Let $x \in A$ and $y \in B$ so that $x + y \in A + B$. Then $x + y \leq \sup A + \sup B$ which implies that $\sup A + \sup B$ is an upper bound for $A + B$. Since $\sup(A + B)$ is the least upper bound this implies that $\sup(A + B) \leq \sup A + \sup B$. Now let $\epsilon \in \mathbb{R}_{>0}$ and let $x \in A$ and $y \in B$ satisfy $\sup A - x < \frac{\epsilon}{2}$ and $\sup B - y < \frac{\epsilon}{2}$. Then

$$\sup A + \sup B - (x + y) < \epsilon.$$

Thus, for any $\epsilon \in \mathbb{R}_{>0}$, there exists $x + y \in A + B$ such that $\sup A + \sup B - (x + y) < \epsilon$. Therefore, $\sup A + \sup B \leq \sup(A + B)$.

(ii) Let $x \in -A$. Then $\sup(-A) \geq x$ or $-\sup(-A) \leq -x$. Thus $-\sup(-A)$ is a lower bound for A and so $\inf(A) \geq -\sup(-A)$. Next let $\epsilon \in \mathbb{R}_{>0}$ and let $x \in -A$ satisfy $x + \epsilon > \sup(-A)$. Then $-x - \epsilon < -\sup(-A)$. Thus, for every $\epsilon \in \mathbb{R}_{>0}$, there exists $y \in A$ such that $y - (-\sup(-A)) < \epsilon$. Thus $-\sup(-A) \geq \inf(A)$, giving this part of the result.

(iii) Let $x \in A$ and note that since $\sup(A) \geq x$, we have $a\sup(A) \geq ax$. Thus $a\sup(A)$ is an upper bound for aA, and so we must have $\sup(aA) \leq a\sup(A)$. Now let $\epsilon \in \mathbb{R}_{>0}$ and let $x \in A$ be such that $x + \frac{\epsilon}{a} > \sup(A)$. Then $ax + \epsilon > a\sup(A)$. Thus, given $\epsilon \in \mathbb{R}_{>0}$ there exists $y \in aA$ such that $a\sup(A) - ax < \epsilon$. Thus $a\sup(A) \leq \sup(aA)$.

(iv) ***missing stuff*** ■

For inf the result is, of course, quite similar. We leave the proof, which mirrors the above proof for sup, to the reader.

2.2.28 Proposition (Properties of inf) *For subsets* $A, B \subseteq \mathbb{R}$ *and for* $a \in \mathbb{R}_{\geq 0}$, *the following statements hold:*

(i) *if* $A + B = \{x + y \mid x \in A,\ y \in B\}$, *then* $\inf(A + B) = \inf(A) + \inf(B)$*;*

(ii) *if* $-A = \{-x \mid x \in A\}$, *then* $\inf(-A) = -\sup(A)$*;*

(iii) *if* $aA = \{ax \mid x \in A\}$, *then* $\inf(aA) = a\inf(A)$*;*

(iv) *if* $I \subseteq \mathbb{R}$ *is an interval, if* $A \subseteq \mathbb{R}$*, if* $f\colon I \to \mathbb{R}$ *is strictly monotonically (see Definition 3.1.27), and if* $f(A) = \{f(x) \mid x \in A\}$, *then* $\inf(f(A)) = f(\inf(A))$.

If $S \subseteq \mathbb{R}$ is a *finite* set, then both $\sup S$ and $\inf S$ are elements of S. In this case we might denote $\max S = \sup S$ and $\min S = \inf S$.

2.2.7 Notes

The Archimedean property of $\mathbb{R}$ seems obvious. The lack of the Archimedean property would mean that there exists t for which $t > N$ for every natural number N. This property is actually possessed by certain fields used in so-called "nonstandard analysis," and we refer the interested reader to **[AR:74]**.

Theorem 2.2.18 is due to **JPGLD:42**, and the proof is a famous use of the "pigeonhole principle." Theorem 2.2.20 is due to **[LK:99]**, and the proof we give is from **[KLK:86]**.

Exercises

2.2.1 Prove the ***Binomial Theorem*** which states that, for $x, y \in \mathbb{R}$ and $k \in \mathbb{Z}_{>0}$,

$$(x + y)^k = \sum_{j=0}^{k} B_{k,j} x^j y^{k-j},$$

where

$$B_{k,j} = \binom{k}{j} \triangleq \frac{k!}{j!(k-j)!}, \qquad j, k \in \mathbb{Z}_{>0},\ j \leq k,$$

are the ***binomial coefficients***, and $k! = 1 \cdot 2 \cdot \cdots \cdot k$ is the ***factorial*** of k. We take the convention that $0! = 1$.

2.2.2 Let $q \in \mathbb{Q} \setminus \{0\}$ and $x \in \mathbb{R} \setminus \mathbb{Q}$. Show the following:

(a) $q + x$ is irrational;

(b) qx is irrational;

(c) $\frac{x}{q}$ is irrational;

(d) $\frac{q}{x}$ is irrational.

2.2.3 Prove Corollary 2.2.16.

2.2.4 Prove Proposition 2.2.10.

2.2.5 Show that the order and absolute value on $\mathbb{R}$ agree with those on $\mathbb{Q}$. That is to say, show the following:

(a) for $q, r \in \mathbb{Q}$, $q < r$ if and only if $i_{\mathbb{Q}}(q) < i_{\mathbb{Q}}(r)$;

(b) for $q \in \mathbb{Q}$, $|q| = |i_{\mathbb{Q}}(q)|$.

(Note that we see clearly here the abuse of notation that follows from using $<$ for both the order on $\mathbb{Z}$ and $\mathbb{Q}$ and from using $|\cdot|$ as the absolute value both on $\mathbb{Z}$ and $\mathbb{Q}$. It is expected that the reader can understand where the notational abuse occurs.)

2.2.6 Do the following:

(a) show that if $x \in \mathbb{R}_{>0}$ satisfies $x < 1$, then $x^k < x$ for each $k \in \mathbb{Z}_{>0}$ satisfying $k \geq 2$;

(b) show that if $x \in \mathbb{R}_{>0}$ satisfies $x > 1$, then $x^k > x$ for each $k \in \mathbb{Z}_{>0}$ satisfying $k \geq 2$.

2.2.7 Show that, for $t, s \in \mathbb{R}$, $||t| - |s|| \leq |t - s|$.

2.2.8 Show that if $s, t \in \mathbb{R}$ satisfy $s < t$, then there exists $q \in \mathbb{Q}$ such that $s < q < t$.

Section 2.3

Sequences in $\mathbb{R}$

In our construction of the real numbers, sequences played a key rôle, inasmuch as Cauchy sequences of rational numbers were integral to our definition of real numbers. In this section we study sequences of real numbers. In particular, in Theorem 2.3.5 we prove the result, absolutely fundamental in analysis, that $\mathbb{R}$ is "complete," meaning that Cauchy sequences of real numbers converge.

Do I need to read this section? If you do not already know the material in this section, then it ought to be read. It is also worth the reader spending some time over the idea that Cauchy sequences of real numbers converge, as compared to rational numbers where this is not the case. The same idea will arise in more abstract settings in Chapter **??**, and so it will pay to understand it well in the simplest case. •

2.3.1 Definitions and properties of sequences

In this section we consider the extension to $\mathbb{R}$ of some of the ideas considered in Section 2.1.2 concerning sequences in $\mathbb{Q}$. As we shall see, it is via sequences, and other equivalent properties, that the nature of the difference between $\mathbb{Q}$ and $\mathbb{R}$ is spelled out quite clearly.

We begin with definitions, generalising in a trivial way the similar definitions for $\mathbb{Q}$.

2.3.1 **Definition (Cauchy sequence, convergent sequence, bounded sequence, monotone sequence)** Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$. The sequence:

(i) is a ***Cauchy sequence*** if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|x_j - x_k| < \epsilon$ for $j,k \geq N$;

(ii) ***converges to*** $\mathbf{s_0}$ if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|x_j - s_0| < \epsilon$ for $j \geq N$;

(iii) ***diverges*** if it does not converge to any element in $\mathbb{R}$;

(iv) is ***bounded above*** if there exists $M \in \mathbb{R}$ such that $x_j < M$ for each $j \in \mathbb{Z}_{>0}$;

(v) is ***bounded below*** if there exists $M \in \mathbb{R}$ such that $x_j > M$ for each $j \in \mathbb{Z}_{>0}$;

(vi) is ***bounded*** if there exists $M \in \mathbb{R}_{>0}$ such that $|x_j| < M$ for each $j \in \mathbb{Z}_{>0}$;

(vii) is ***monotonically increasing*** if $x_{j+1} \geq x_j$ for $j \in \mathbb{Z}_{>0}$;

(viii) is ***strictly monotonically increasing*** if $x_{j+1} > x_j$ for $j \in \mathbb{Z}_{>0}$;

(ix) is ***monotonically decreasing*** if $x_{j+1} \leq x_j$ for $j \in \mathbb{Z}_{>0}$;

(x) is ***strictly monotonically decreasing*** if $x_{j+1} < x_j$ for $j \in \mathbb{Z}_{>0}$;

(xi) is ***constant*** if $x_j = x_1$ for every $j \in \mathbb{Z}_{>0}$;

(xii) is ***eventually constant*** if there exists $N \in \mathbb{Z}_{>0}$ such that $x_j = x_N$ for every $j \geq N$. •

Associated with the notion of convergence is the notion of a limit. We also, for convenience, wish to allow sequences with infinite limits. This makes for some rather subtle use of language, so the reader should pay attention to this.

2.3.2 Definition (Limit of a sequence) Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence.

(i) If $(x_j)_{j\in\mathbb{Z}_{>0}}$ converges to s_0, then the sequence has s_0 as a ***limit***, and we write $\lim_{j\to\infty} x_j = s_0$.

(ii) If, for every $M \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $x_j > M$ (resp. $x_k < -M$) for $j \geq N$, then the sequence ***diverges to*** $\boldsymbol{\infty}$ (resp. ***diverges to*** $\boldsymbol{-\infty}$), and we write $\lim_{j\to\infty} x_j = \infty$ (resp. $\lim_{j\to\infty} x_j = -\infty$);

(iii) If $\lim_{j\to\infty} x_j \in \mathbb{R}$, then the limit of the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ ***exists***.

(iv) If the limit of the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ does not exist, does not diverge to ∞, or does not diverge to $-\infty$, then the sequence is ***oscillatory***. •

The reader can prove in Exercise 2.3.1 that limits, if they exist, are unique.

That convergent sequences are Cauchy, and that Cauchy sequences are bounded follows in exactly the same manner as the analogous results, stated as Propositions 2.1.13 and 2.1.14, for $\mathbb{Q}$. Let us state the results here for reference.

2.3.3 Proposition (Convergent sequences are Cauchy) *If a sequence* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* x_0*, then it is a Cauchy sequence.*

2.3.4 Proposition (Cauchy sequences are bounded) *If* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence in* $\mathbb{R}$ *then it is bounded.*

Moreover, what is true for $\mathbb{R}$, and that is not true for $\mathbb{Q}$, is that every Cauchy sequence converges.

2.3.5 Theorem (Cauchy sequences in $\mathbb{R}$ converge) *If* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a Cauchy sequence in* $\mathbb{R}$ *then there exists* $\mathrm{s}_0 \in \mathbb{R}$ *such that* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* s_0*.*

Proof For $j \in \mathbb{Z}_{>0}$ choose $q_j \in \mathbb{Q}_{>0}$ such that $|x_j - q_j| < \frac{1}{j}$, this being possible by Proposition 2.2.15. For $\epsilon \in \mathbb{R}_{>0}$ let $N_1 \in \mathbb{Z}_{>0}$ satisfy $|x_j - x_k| < \frac{\epsilon}{2}$ for $j, k \geq N_1$. By Proposition 2.2.13 let $N_2 \in \mathbb{Z}_{>0}$ satisfy $N_2 \cdot 1 > 4\epsilon^{-1}$, and let N be the larger of N_1 and N_2. Then, for $j, k \geq N$, we have

$$|q_j - q_k| = |q_j - x_j + x_j - x_k + x_k - q_k| \leq |x_j - q_j| + |x_j - x_k| + |x_k - q_k| < \tfrac{1}{j} + \tfrac{\epsilon}{2} + \tfrac{1}{k} < \epsilon.$$

Thus $(q_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence, and so we define $s_0 = [(q_j)_{j\in\mathbb{Z}_{>0}}]$.

Now we show that $(q_j)_{j\in\mathbb{Z}_{>0}}$ converges to s_0. Let $\epsilon \in \mathbb{R}_{>0}$ and take $N \in \mathbb{Z}_{>0}$ such that $|q_j - q_k| < \frac{\epsilon}{2}$, $j, k \geq N$, and rewrite this as

$$\tfrac{\epsilon}{2} < q_j - q_k + \epsilon, \quad \tfrac{\epsilon}{2} < -q_k + q_k + \epsilon, \qquad j, k \geq N. \tag{2.4}$$

For $j_0 \geq N$ consider the sequence $(q_j - q_{j_0} + \epsilon)_{j\in\mathbb{Z}_{>0}}$. This is a Cauchy sequence by Proposition 2.2.1. Moreover, by Proposition 2.2.6, $[(q_j - q_{j_0} + \epsilon)_{j\in\mathbb{Z}_{>0}}] > 0$, using the first of the inequalities in (2.4). Thus we have $s_0 - q_{j_0} + \epsilon > 0$, or

$$-\epsilon < s_0 - q_{j_0}, \qquad j_0 \geq N.$$

Arguing similarly, but using the second of the inequalities (2.4), we determine that

$$s_0 - q_{j_0} < \epsilon, \qquad j_0 \geq N.$$

This gives $|s_0 - q_j| < \epsilon$ for $j \geq N$, so showing that $(q_j)_{j\in\mathbb{Z}_{>0}}$ converges to s_0.

Finally, we show that $(x_j)_{j\in\mathbb{Z}_{>0}}$ converges to s_0. Let $\epsilon \in \mathbb{R}_{>0}$ and take $N_1 \in \mathbb{Z}_{>0}$ such that $|s_0 - q_j| < \frac{\epsilon}{2}$ for $j \geq N_1$. Also choose $N_2 \in \mathbb{Z}_{>0}$ such that $N_2 \cdot 1 > 2\epsilon^{-1}$ by Proposition 2.2.13. If N is the larger of N_1 and N_2, then we have

$$|s_0 - x_j| = |s_0 - q_j + q_j - x_j| \leq |s_0 - q_j| + |q_j - x_j| < \tfrac{\epsilon}{2} + \tfrac{1}{j} < \epsilon,$$

for $j \geq N$, so giving the result. ■

2.3.6 Remark (Completeness of $\mathbb{R}$) The property of $\mathbb{R}$ that Cauchy sequences are convergent gives, in the more general setting of Section **??**, $\mathbb{R}$ the property of being ***complete***. Completeness is an extremely important concept in analysis. We shall say some words about this in Section 6.3.2; for now let us just say that the subject of calculus would not exist, but for the completeness of $\mathbb{R}$. •

2.3.2 Some properties equivalent to the completeness of $\mathbb{R}$

Using the fact that Cauchy sequences converge, it is easy to prove two other important features of $\mathbb{R}$, both of which seem obvious intuitively.

2.3.7 Theorem (Bounded subsets of $\mathbb{R}$ have a least upper bound) *If* $S \subseteq \mathbb{R}$ *is nonempty and possesses an upper bound with respect to the standard total order* $\leq$*, then* S *possesses a least upper bound with respect to the same total order.*

Proof Since S has an upper bound, there exists $y \in \mathbb{R}$ such that $x \leq y$ for all $x \in S$. Now choose some $x \in S$. We then define two sequences $(x_j)_{j\in\mathbb{Z}_{>0}}$ and $(y_j)_{j\in\mathbb{Z}_{>0}}$ recursively as follows:

1. define $x_1 = x$ and $y_1 = y$;
2. suppose that x_j and y_j have been defined;
3. if there exists $z \in S$ with $\frac{1}{2}(x_j + y_j) < z \leq y_j$, take $x_{j+1} = z$ and $y_{j+1} = y_j$;
4. if there is no $z \in S$ with $\frac{1}{2}(x_j + y_j) < z \leq y_j$, take $x_{j+1} = x_j$ and $y_{j+1} = \frac{1}{2}(x_j + y_j)$.

A lemma characterises these sequences.

1 Lemma *The sequences* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *and* $(y_j)_{j\in\mathbb{Z}_{>0}}$ *have the following properties:*

(i) $x_j \in S$ *for* $j \in \mathbb{Z}_{>0}$*;*

(ii) $x_{j+1} \geq x_j$ *for* $j \in \mathbb{Z}_{>0}$*;*

(iii) y_j *is an upper bound for* S *for* $j \in \mathbb{Z}_{>0}$*;*

(iv) $y_{j+1} \leq y_j$ *for* $j \in \mathbb{Z}_{>0}$*;*

(v) $0 \leq y_j - x_j \leq \frac{1}{2^j}(y - x)$ *for* $j \in \mathbb{Z}_{>0}$.

Proof We prove the result by induction on j. The result is obviously true for $= 0$. Now suppose the result true for $j \in \{1, \ldots, k\}$.

First take the case where there exists $z \in S$ with $\frac{1}{2}(x_k + y_k) < z \leq y_k$, so that $x_{k+1} = z$ and $y_{k+1} = y_k$. Clearly $x_{k+1} \in S$ and $y_{k+1} \geq y_k$. Since $y_k \geq x_k$ by the induction hypotheses, $\frac{1}{2}(x_k + y_k) \geq x_k$ giving $x_{k+1} = z \geq x_k$. By the induction hypotheses, y_{k+1} is an upper bound for S. By definition of x_{k+1} and y_{k+1},

$$y_{k+1} - x_{k+1} = y_k - z \geq 0$$

and

$$y_{k+1} - x_{k+1} = y_k - z = y_k - \tfrac{1}{2}(y_k - x_k) = \tfrac{1}{2}(y_k - x_k),$$

giving $y_{k+1} - x_{k+1} \leq \frac{1}{2^{k+1}}(y - x)$ by the induction hypotheses.

Now we take the case where there is no $z \in S$ with $\frac{1}{2}(x_j + y_j) < z \leq y_j$, so that $x_{k+1} = x_k$ and $y_{k+1} = \frac{1}{2}(x_k + y_k)$. Clearly $x_{k+1} \geq x_k$ and $x_{k+1} \in S$. If y_{k+1} were not an upper bound for S, then there exists $a \in S$ such that $a > y_{k+1}$. By the induction hypotheses, y_k is an upper bound for S so $a \leq y_k$. But this means that $\frac{1}{2}(y_k + x_k) < a \leq y_k$, contradicting our assumption concerning the nonexistence of $z \in S$ with $\frac{1}{2}(x_j + y_j) < z \leq y_j$. Thus y_{k+1} is an upper bound for S. Since $x_k \leq y_k$ by the induction hypotheses,

$$y_{k+1} = \tfrac{1}{2}(y_k + x_k) \leq y_k.$$

Also

$$y_{k+1} - x_{k+1} = \tfrac{1}{2}(y_k - x_k)$$

by the induction hypotheses. This completes the proof. ▼

The following lemma records a useful fact about the sequences $(x_j)_{j\in\mathbb{Z}_{>0}}$ and $(y_j)_{j\in\mathbb{Z}_{>0}}$.

2 Lemma *Let* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\mathrm{y_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be sequences in* $\mathbb{R}$ *satisfying:*

(i) $\mathrm{x_{j+1} \geq x_j}$, $\mathrm{j} \in \mathbb{Z}_{>0}$;

(ii) $\mathrm{y_{j+1} \leq y_j}$, $\mathrm{j} \in \mathbb{Z}_{>0}$;

(iii) the sequence $(\mathrm{y_j - x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* 0.

Then $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\mathrm{y_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converge, and converge to the same limit.*

Proof First we claim that $x_j \leq y_k$ for all $j, k \in \mathbb{Z}_{>0}$. Indeed, suppose not. Then there exists $j, k \in \mathbb{Z}_{>0}$ such that $x_j > y_k$. If N is the larger of j and k, then we have $y_N \leq y_k < x_j \leq x_N$. This implies that

$$x_m - y_m \geq x_j - y_m \geq x_j - y_k > 0, \qquad m \geq N,$$

which contradicts the fact that $(y_j - x_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero.

Now, for $\epsilon \in \mathbb{R}_{>0}$ let $N \in \mathbb{Z}_{>0}$ satisfy $|y_j - x_j| < \epsilon$ for $j \geq N$, or, simply, $y_j - x_j < \epsilon$ for $j \geq N$. Now let $j, k \geq N$, and suppose that $j \geq k$. Then

$$0 \leq x_j - x_k \leq x_j - y_k < \epsilon.$$

Similarly, if $j \leq k$ we have $0 \leq x_k - x_j < \epsilon$. In other words, $|x_j - x_k| < \epsilon$ for $j, k \geq N$. Thus $(x_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. In like manner one shows that $(y_j)_{j\in\mathbb{Z}_{>0}}$ is also a Cauchy sequence. Therefore, by Theorem 2.3.5, these sequences converge, and let us denote their limits by s_0 and t_0, respectively. However, since $(x_j)_{j\in\mathbb{Z}_{>0}}$ and $(y_j)_{j\in\mathbb{Z}_{>0}}$ are equivalent Cauchy sequences in the sense of Definition 2.1.16, it follows that $s_0 = t_0$. ▼

Using Lemma 1 we easily verify that the sequences $(x_j)_{j\in\mathbb{Z}_{>0}}$ and $(y_j)_{j\in\mathbb{Z}_{>0}}$ satisfy the hypotheses of Lemma 2. Therefore these sequences converge to a common limit, which we denote by s. We claim that s is a least upper bound for S. First we show that it is an upper bound. Suppose that there is $x \in S$ such that $x > s$ and define $\epsilon = x - s$. Since $(y_j)_{j\in\mathbb{Z}_{>0}}$ converges to s, there exists $N \in \mathbb{Z}_{>0}$ such that $|s - y_j| < \epsilon$ for $j \geq N$. Then, for $j \geq N$,

$$y_j - s < \epsilon = x - s,$$

implying that $y_j < x$, and so contradicting Lemma 1.

Finally, we need to show that s is a least upper bound. To see this, let b be an upper bound for S and suppose that $b < s$. Define $\epsilon = s - b$, and choose $N \in \mathbb{Z}_{>0}$ such that $|s - x_j| < \epsilon$ for $j \geq N$. Then

$$s - x_j < \epsilon = s - b,$$

implying that $b < x_j$ for $j \geq N$. This contradicts the fact, from Lemma 1, that $x_j \in S$ and that b is an upper bound for S. ■

As we shall explain more fully in Aside 2.3.9, the least upper bound property of the real numbers as stated in the preceding theorem is actually *equivalent* to the completeness of $\mathbb{R}$. In fact, the least upper bound property forms the basis for an alternative definition of the real numbers using ***Dedekind cuts***.[4] Here the idea is that one defines a real number as being a splitting of the rational numbers into two halves, one corresponding to the rational numbers less than the real number one is defining, and the other corresponding to the rational numbers greater than the real number one is defining. Historically, Dedekind cuts provided the first rigorous construction of the real numbers. We refer to Section 2.3.9 for further discussion. We also comment, as we discuss in Aside 2.3.9, that any construction of the real numbers with the property of completeness, or an equivalent, will produce something that is "essentially" the real numbers as we have defined them.

Another consequence of Theorem 2.3.5 is the following.

2.3.8 Theorem (Bounded, monotonically increasing sequences in $\mathbb{R}$ converge) *If* $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ *is a bounded, monotonically increasing sequence in* $\mathbb{R}$*, then it converges.*

Proof The subset $(x_j)_{j\in\mathbb{Z}_{>0}}$ of $\mathbb{R}$ has an upper bound, since it is bounded. By Theorem 2.3.7 let b be the least upper bound for this set. We claim that $(x_j)_{j\in\mathbb{Z}_{>0}}$ converges to b. Indeed, let $\epsilon \in \mathbb{R}_{>0}$. We claim that there exists some $N \in \mathbb{Z}_{>0}$ such that $b - x_N < \epsilon$ since b is a least upper bound. Indeed, if there is no such N, then $b \geq x_j + \epsilon$ for all $j \in \mathbb{Z}_{>0}$ and so $b - \frac{\epsilon}{2}$ is an upper bound for $(x_j)_{j\in\mathbb{Z}_{>0}}$ that is smaller than b. Now, with N chosen so that $b - x_N < \epsilon$, the fact that $(x_j)_{j\in\mathbb{Z}_{>0}}$ is monotonically increasing implies that $|b - x_j| < \epsilon$ for $j \geq N$, as desired. ■

It turns out that Theorems 2.3.5, 2.3.7, and 2.3.8 are equivalent. But to make sense of this requires one to step outside the concrete representation we have given for the real numbers to a more axiomatic one. This can be skipped, so we present it as an aside.

[4] After Julius Wihelm Richard Dedekind (1831–1916), the German mathematician, did work in the areas of analysis, ring theory, and set theory. His rigorous mathematical style has had a strong influence on modern mathematical presentation.

2.3.9 Aside (Complete ordered fields) An ***ordered field*** is a field $\mathbb{F}$ (see Definition 4.2.1 for the definition of a field) equipped with a total order satisfying the conditions

1. if $x < y$ then $x + z < y + z$ for $x, y, z \in \mathbb{F}$ and
2. if $0 < x, y$ then $0 < x \cdot y$.

Note that in an ordered field one can define the absolute value exactly as we have done for $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. There are many examples of ordered fields, of which $\mathbb{Q}$ and $\mathbb{R}$ are two that we have seen. However, if one adds to the conditions for an ordered field an additional condition, then this turns out to essentially uniquely specify the set of real numbers. (We say "essentially" since the uniqueness is up to a bijection that preserves the field structure as well as the order.) This additional structure comes in various forms, of which three are as stated in Theorems 2.3.5, 2.3.7, and 2.3.8. To be precise, we have the following theorem.

Theorem *If $\mathbb{F}$ is an ordered field, then the following statements are equivalent:*

(i) every Cauchy sequence converges;

(ii) each set possessing an upper bound possesses a least upper bound;

(iii) each bounded, monotonically increasing sequence converges.

We have almost proved this theorem with our arguments above. To see this, note that in the proof of Theorem 2.3.7 we use the fact that Cauchy sequences converge. Moreover, the argument can easily be adapted from the special case of $\mathbb{R}$ to a general ordered field. This gives the implication (i) $\implies$ (ii) in the theorem above. In like manner, the proof of Theorem 2.3.8 gives the implication (ii) $\implies$ (iii), since the proof is again easily seen to be valid for a general ordered field. The argument for the implication (iii) $\implies$ (i) is outlined in Exercise 2.3.5. An ordered field satisfying any one of the three equivalent conditions (i), (ii), and (iii) is called a ***complete ordered field***. Thus there is essentially only one complete ordered field, and it is $\mathbb{R}$. ♠

2.3.3 Tests for convergence of sequences

There is generally no algorithmic way, other than checking the definition, to ascertain when a sequence converges. However, there are a few simple results that are often useful, and here we state some of these.

2.3.10 Proposition (Squeezing Principle) *Let $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$, $(\mathrm{y}_j)_{j\in\mathbb{Z}_{>0}}$, and $(\mathrm{z}_j)_{j\in\mathbb{Z}_{>0}}$ be sequences in $\mathbb{R}$ satisfying*

(i) $\mathrm{x}_j \le \mathrm{z}_j \le \mathrm{y}_j$ for all $j \in \mathbb{Z}_{>0}$ and

(ii) $\lim_{j\to\infty} \mathrm{x}_j = \lim_{j\to\infty} \mathrm{y}_j = \alpha$.

Then $\lim_{j\to\infty} \mathrm{z}_j = \alpha$.

Proof Let $\epsilon \in \mathbb{R}_{>0}$ and let $N_1, N_2 \in \mathbb{Z}_{>0}$ have the property that $|x_j - \alpha| < \frac{\epsilon}{3}$ for $j \ge N_1$ and $|y_j - \alpha| < \frac{\epsilon}{3}$. Then, for $j \ge \max\{N_1, N_2\}$,

$$|x_j - y_j| = |x_j - \alpha + \alpha - y_j| \le |x_j - \alpha| + |y_j - \alpha| < \tfrac{2\epsilon}{3},$$

using the triangle inequality. Then, for $j \geq \max\{N_1, N_2\}$, we have

$$|z_j - \alpha| = |z_j - x_j + x_j - \alpha| \leq |z_j - x_j| + |x_j - \alpha| \leq |y_j - x_j| + |x_j - \alpha| = \epsilon,$$

again using the triangle inequality. ■

The next test for convergence of a series is sometimes useful.

2.3.11 Proposition (Ratio Test for sequences) *Let* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathbb{R}$ *for which* $\lim_{\mathrm{j}\to\infty}\left|\frac{\mathrm{x_{j+1}}}{\mathrm{x_j}}\right| = \alpha$*. If* $\alpha < 1$ *then the sequence* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* 0*, and if* $\alpha > 1$ *then the sequence* $(\mathrm{x_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *diverges.*

Proof For $\alpha < 1$, define $\beta = \frac{1}{2}(\alpha + 1)$. Then $\alpha < \beta < 1$. Now take $N \in \mathbb{Z}_{>0}$ such that

$$\left|\left|\frac{x_{j+1}}{x_j}\right| - \alpha\right| < \tfrac{1}{2}(1 - \alpha), \qquad j > N.$$

This implies that

$$\left|\frac{x_{j+1}}{x_j}\right| < \beta.$$

Now, for $j > N$,

$$|x_j| < \beta|x_{j-1}| < \beta^2|x_{j-1}| < \cdots < \beta^{j-N}|x_N|.$$

Clearly the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ converges to 0 if and only if the sequence obtained by replacing the first N terms by 0 also converges to 0. If this latter sequence is denoted by $(y_j)_{j\in\mathbb{Z}_{>0}}$, then we have

$$0 \leq y_j \leq \frac{|x_N|}{\beta^N}\beta^j.$$

The sequence $(\frac{|x_N|}{\beta^N}\beta^j)_{j\in\mathbb{Z}_{>0}}$ converges to 0 since $\beta < 1$, and so this part of the result follows from the Squeezing Principle.

For $\alpha > 1$, there exists $N \in \mathbb{Z}_{>0}$ such that, for all $j \geq N$, $x_j \neq 0$. Consider the sequence $(y_j)_{j\in\mathbb{Z}_{>0}}$ which is 0 for the first N terms, and satisfies $y_j = x_j^{-1}$ for the remaining terms. We then have $\left|\frac{y_{j+1}}{y_j}\right| < \alpha^{-1} < 1$, and so, from the first part of the proof, the sequence $(y_j)_{j\in\mathbb{Z}_{>0}}$ converges to 0. Thus the sequence $(|y_j|)_{j\in\mathbb{Z}_{>0}}$ converges to ∞, which prohibits the sequence $(y_j)_{j\in\mathbb{Z}_{>0}}$ from converging. ■

In Exercise 2.3.3 the reader can explore the various possibilities for the ratio test when $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$.

2.3.4 lim sup and lim inf

Recall from Section 2.2.6 the notions of sup and inf for subsets of $\mathbb{R}$. Associated with the least upper bound and greatest lower bound properties of $\mathbb{R}$ is a useful notion that weakens the usual idea of convergence. In order for us to make a sensible definition, we first prove a simple result.

2.3.12 Proposition (Existence of lim sup and lim inf) *For any sequence* $(\mathrm{x}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *in* $\mathbb{R}$*, the limits*

$$\lim_{\mathrm{N}\to\infty}\big(\sup\{\mathrm{x}_\mathrm{j} \mid \mathrm{j}\geq \mathrm{N}\}\big),\quad \lim_{\mathrm{N}\to\infty}\big(\inf\{\mathrm{x}_\mathrm{j} \mid \mathrm{j}\geq \mathrm{N}\}\big)$$

exist, diverge to ∞*, or diverge to* $-\infty$*.*

Proof Note that the sequences $(\sup\{x_j \mid j\geq N\})_{N\in\mathbb{Z}_{>0}}$ and $(\inf\{x_j \mid j\geq N\})_{N\in\mathbb{Z}_{>0}}$ in $\overline{\mathbb{R}}$ are monotonically decreasing and monotonically increasing, respectively, with respect to the natural order on $\overline{\mathbb{R}}$. Moreover, note that a monotonically increasing sequence in $\overline{\mathbb{R}}$ is either bounded by some element of $\mathbb{R}$, or it is not. If the sequence is upper bounded by some element of $\mathbb{R}$, then by Theorem 2.3.8 it either converges or is the sequence $(-\infty)_{j\in\mathbb{Z}_{>0}}$. If it is not bounded by some element in $\mathbb{R}$, then either it diverges to ∞, or it is the sequence $(\infty)_{j\in\mathbb{Z}_{>0}}$ (this second case cannot arise in the specific case of the monotonically increasing sequence $(\sup\{x_j \mid j\geq N\})_{N\in\mathbb{Z}_{>0}}$. In all cases, the limit $\lim_{N\to\infty}\big(\sup\{x_j \mid j\geq N\}\big)$ exists or diverges to ∞. A similar argument for holds for $\lim_{N\to\infty}\big(\inf\{x_j \mid j\geq N\}\big)$. ■

2.3.13 Definition (lim sup and lim inf) For a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ denote

$$\limsup_{j\to\infty} x_j = \lim_{N\to\infty}\big(\sup\{x_j \mid j\geq N\}\big),$$
$$\liminf_{j\to\infty} x_j = \lim_{N\to\infty}\big(\inf\{x_j \mid j\geq N\}\big).$$
•

Before we get to characterising lim sup and lim inf, we give some examples to illustrate all the cases that can arise.

2.3.14 Examples (lim sup and lim inf)

1. Consider the sequence $(x_j = (-1)^j)_{j\in\mathbb{Z}_{>0}}$. Here we have $\limsup_{j\to\infty} x_j = 1$ and $\liminf_{j\to\infty} x_j = -1$.
2. Consider the sequence $(x_j = j)_{j\in\mathbb{Z}_{>0}}$. Here $\limsup_{j\to\infty} x_j = \liminf_{j\to\infty} = \infty$.
3. Consider the sequence $(x_j = -j)_{j\in\mathbb{Z}_{>0}}$. Here $\limsup_{j\to\infty} x_j = \liminf_{j\to\infty} = -\infty$.
4. Define
$$x_j = \begin{cases} j, & j \text{ even},\\ 0, & j \text{ odd}.\end{cases}$$
We then have $\limsup_{j\to\infty} x_j = \infty$ and $\liminf_{j\to\infty} x_j = 0$.
5. Define
$$x_j = \begin{cases} -j, & j \text{ even},\\ 0, & j \text{ odd}.\end{cases}$$
We then have $\limsup_{j\to\infty} x_j = 0$ and $\liminf_{j\to\infty} = -\infty$.
6. Define
$$x_j = \begin{cases} j, & j \text{ even},\\ -j, & j \text{ odd}.\end{cases}$$
We then have $\limsup_{j\to\infty} x_j = \infty$ and $\liminf_{j\to\infty} = -\infty$. •

There are many ways to characterise lim sup and lim inf, and we shall indicate but a few of these.

2.3.15 Proposition (Characterisation of lim sup) *For a sequence $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ and $\alpha \in \mathbb{R}$, the following statements are equivalent:*

(i) $\alpha = \limsup_{j\to\infty} \mathrm{x}_j$*;*

(ii) $\alpha = \inf\{\sup\{\mathrm{x}_j \mid j \geq k\} \mid k \in \mathbb{Z}_{>0}\}$*;*

(iii) for each $\epsilon \in \mathbb{R}_{>0}$ *the following statements hold:*

(a) there exists $\mathrm{N} \in \mathbb{Z}_{>0}$ *such that* $\mathrm{x}_j < \alpha + \epsilon$ *for all* $j \geq \mathrm{N}$*;*

(b) for an infinite number of $j \in \mathbb{Z}_{>0}$ *it holds that* $\mathrm{x}_j > \alpha - \epsilon$*.*

Proof (i) $\iff$ (ii) Let $y_k = \sup\{x_j \mid j \geq k\}$ and note that the sequence $(y_k)_{k\in\mathbb{Z}_{>0}}$ is monotonically decreasing. Therefore, the sequence $(y_k)_{k\in\mathbb{Z}_{>0}}$ converges if and only if it is lower bounded. Moreover, if it converges, it converges to $\inf(y_k)_{k\in\mathbb{Z}_{>0}}$. Putting this all together gives the desired implications.

(i) $\Longrightarrow$ (iii) Let y_k be as in the preceding part of the proof. Since $\lim_{k\to\infty} y_k = \alpha$, for each $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that $|y_k - \alpha| < \epsilon$ for $k \geq N$. In particular, $y_N < \alpha + \epsilon$. Therefore, $x_j < \alpha + \epsilon$ for all $j \geq N$, so (iii a) holds. We also claim that, for every $\epsilon \in \mathbb{R}_{>0}$ and for every $N \in \mathbb{Z}_{>0}$, there exists $j \geq N$ such that $x_j > y_N - \epsilon$. Indeed, if $x_j \leq y_N - \epsilon$ for every $j \geq N$, then this contradicts the definition of y_N. Since $y_N \geq \alpha$ we have $x_j > y_N - \epsilon \geq \alpha - \epsilon$ for some j. Since N is arbitrary, (iii b) holds.

(iii) $\Longrightarrow$ (i) Condition (iii a) means that there exists $N \in \mathbb{Z}_{>0}$ such that $y_k < \alpha + \epsilon$ for all $k \geq N$. Condition (iii b) implies that $y_k > \alpha - \epsilon$ for all $k \in \mathbb{Z}_{>0}$. Combining these conclusions shows that $\lim_{k\to\infty} y_k = \alpha$, as desired. ■

The corresponding result for lim inf is the following. The proof follows in the same manner as the result for lim sup.

2.3.16 Proposition (Characterisation of lim inf) *For a sequence $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ and $\alpha \in \mathbb{R}$, the following statements are equivalent:*

(i) $\alpha = \liminf_{j\to\infty} \mathrm{x}_j$*;*

(ii) $\alpha = \sup\{\inf\{\mathrm{x}_j \mid j \geq k\} \mid k \in \mathbb{Z}_{>0}\}$*;*

(iii) for each $\epsilon \in \mathbb{R}_{>0}$ *the following statements hold:*

(a) there exists $\mathrm{N} \in \mathbb{Z}_{>0}$ *such that* $\mathrm{x}_j > \alpha - \epsilon$ *for all* $j \geq \mathrm{N}$*;*

(b) for an infinite number of $j \in \mathbb{Z}_{>0}$ *it holds that* $\mathrm{x}_j < \alpha + \epsilon$*.*

Finally, we characterise the relationship between lim sup, lim inf, and lim.

2.3.17 Proposition (Relationship between lim sup, lim inf, and lim) *For a sequence $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ and $\mathrm{s}_0 \in \mathbb{R}$, the following statements are equivalent:*

(i) $\lim_{j\to\infty} \mathrm{x}_j = \mathrm{s}_0$*;*

(ii) $\limsup_{j\to\infty} \mathrm{x}_j = \liminf_{j\to\infty} \mathrm{x}_j = \mathrm{s}_0$*.*

Proof (i) $\Longrightarrow$ (ii) Let $\epsilon \in \mathbb{R}_{>0}$ and take $N \in \mathbb{Z}_{>0}$ such that $|x_j - s_0| < \epsilon$ for all $j \geq N$. Then $x_j < s_0 + \epsilon$ and $x_j > s_0 - \epsilon$ for all $j \geq N$. The current implication now follows from Propositions 2.3.15 and 2.3.16.

(ii) $\Longrightarrow$ (i) Let $\epsilon \in \mathbb{R}_{>0}$. By Propositions 2.3.15 and 2.3.16 there exists $N_1, N_2 \in \mathbb{Z}_{>0}$ such that $x_j - s_0 < \epsilon$ for $j \geq N_1$ and $s_0 - x_j < \epsilon$ for $j \geq N_2$. Thus $|x_j - s_0| < \epsilon$ for $j \geq \max\{N_1, N_2\}$, giving this implication. ■

2.3.5 Multiple sequences

It will be sometimes useful for us to be able to consider sequences indexed, not by a single index, but by multiple indices. We consider the case here of two indices, and extensions to more indices are done by induction.

2.3.18 Definition (Double sequence) A ***double sequence*** in $\mathbb{R}$ is a family of elements of $\mathbb{R}$ indexed by $\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$. We denote a double sequence by $(x_{jk})_{j,k\in\mathbb{Z}_{>0}}$, where x_{jk} is the image of $(j,k) \in \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ in $\mathbb{R}$. •

It is not *a priori* obvious what it might mean for a double sequence to converge, so we should carefully say what this means.

2.3.19 Definition (Convergence of double sequences) Let $s_0 \in \mathbb{R}$. A double sequence $(x_{jk})_{j,k\in\mathbb{Z}_{>0}}$:

(i) ***converges to*** $\mathbf{s_0}$, and we write $\lim_{j,k\to\infty} x_{jk} = s_0$, if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|s_0 - x_{jk}| < \epsilon$ for $j,k \geq N$;

(ii) has s_0 as a ***limit*** if it converges to s_0.

(iii) is ***convergent*** if it converges to some member of $\mathbb{R}$;

(iv) ***diverges*** if it does not converge;

(v) ***diverges to*** $\boldsymbol{\infty}$ (resp. ***diverges to*** $\boldsymbol{-\infty}$), and we write $\lim_{j,k\to\infty} x_{jk} = \infty$ (resp. $\lim_{j,k\to\infty} x_{jk} = -\infty$) if, for each $M \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $x_{jk} > M$ (resp. $x_{jk} < -M$) for $j,k \geq N$;

(vi) has a limit that ***exists*** if $\lim_{j,k\to\infty} x_{jk} \in \mathbb{R}$;

(vii) is ***oscillatory*** if the limit of the sequence does not exist, does not diverge to ∞, or does not diverge to $-\infty$. •

Note that the definition of convergence requires that one check both indices at the same time. Indeed, if one thinks, as it is useful to do, of a double sequence as assigning a real number to each point in an infinite grid defined by the set $\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$, convergence means that the values on the grid can be made arbitrarily small outside a sufficiently large square (see Figure 2.2). It is useful, however, to have means of computing limits of double sequences by computing limits of sequences in the usual sense. Our next results are devoted to this.

2.3.20 Proposition (Computation of limits of double sequences I) *Suppose that for the double sequence* $(\mathrm{x_{jk}})_{\mathrm{j,k}\in\mathbb{Z}_{>0}}$ *it holds that*

(i) the double sequence is convergent and

(ii) for each $\mathrm{j} \in \mathbb{Z}_{>0}$*, the limit* $\lim_{\mathrm{k}\to\infty} \mathrm{x_{jk}}$ *exists.*

Then the limit $\lim_{\mathrm{j}\to\infty}(\lim_{\mathrm{k}\to\infty} \mathrm{x_{jk}})$ *exists and is equal to* $\lim_{\mathrm{j,k}\to\infty} \mathrm{x_{jk}}$.

Proof Let $s_0 = \lim_{j,k\to\infty} x_{jk}$ and denote $s_j = \lim_{k\to\infty} x_{jk}$, $j \in \mathbb{Z}_{>0}$. For $\epsilon \in \mathbb{R}_{>0}$ take $N \in \mathbb{Z}_{>0}$ such that $|x_{jk} - s_0| < \frac{\epsilon}{2}$ for $j,k \geq N$. Also take $N_j \in \mathbb{Z}_{>0}$ such that $|x_{jk} - s_j| < \frac{\epsilon}{2}$ for $k \geq N_j$. Next take $j \geq N$ and let $k \geq \max\{N, N_j\}$. We then have

$$|s_j - s_0| = |s_j - x_{jk} + x_{jk} - s_0| \leq |s_j - x_{jk}| + |x_{jk} - s_0| < \epsilon,$$

using the triangle inequality. ■

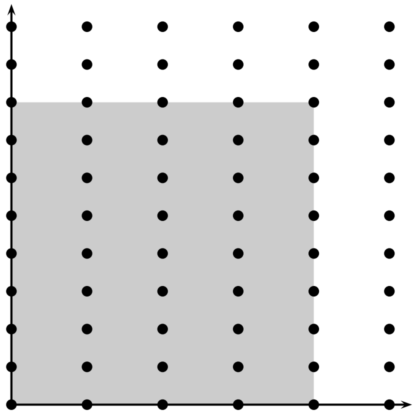


Figure 2.2 Convergence of a double sequence: by choosing the square large enough, the values at the unshaded grid points can be arbitrarily close to the limit

2.3.21 Proposition (Computation of limits of double sequences II) *Suppose that for the double sequence* $(\mathrm{x}_{\mathrm{jk}})_{\mathrm{j,k}\in\mathbb{Z}_{>0}}$ *it holds that*

(i) the double sequence is convergent,

(ii) for each $\mathrm{j}\in\mathbb{Z}_{>0}$*, the limit* $\lim_{\mathrm{k}\to\infty}\mathrm{x}_{\mathrm{jk}}$ *exists, and*

(iii) for each $\mathrm{k}\in\mathbb{Z}_{>0}$*, the limit* $\lim_{\mathrm{j}\to\infty}\mathrm{x}_{\mathrm{jk}}$ *exists.*

Then the limits $\lim_{\mathrm{j}\to\infty}(\lim_{\mathrm{k}\to\infty}\mathrm{x}_{\mathrm{jk}})$ *and* $\lim_{\mathrm{k}\to\infty}(\lim_{\mathrm{j}\to\infty}\mathrm{x}_{\mathrm{jk}})$ *exist and are equal to* $\lim_{\mathrm{j,k}\to\infty}\mathrm{x}_{\mathrm{jk}}$*.*

Proof This follows from two applications of Proposition 2.3.20. ■

Let us give some examples that illustrate the idea of convergence of a double sequence.

2.3.22 Examples (Double sequences)

1. It is easy to check that the double sequence $(\frac{1}{j+k})_{j,k\in\mathbb{Z}_{>0}}$ converges to 0. Indeed, for $\epsilon\in\mathbb{R}_{>0}$, if we take $N\in\mathbb{Z}_{>0}$ such that $\frac{1}{2N}<\epsilon$, it follows that $\frac{1}{j+k}<\epsilon$ for $j,k\geq N$.

2. The double sequence $(\frac{j}{j+k})_{j,k\in\mathbb{Z}_{>0}}$ does not converge. To see this we should find $\epsilon\in\mathbb{R}_{>0}$such that, for any $N\in\mathbb{Z}_{>0}$, there exists $j,k\geq N$ for which $\frac{j}{j+k}\geq\epsilon$. Take $\epsilon=\frac{1}{2}$ and let $N\in\mathbb{Z}_{>0}$. Then, if $j,k\geq N$ satisfy $j\geq 2k$, we have $\frac{j}{j+k}\geq\epsilon$.

   Note that for this sequence, the limits $\lim_{j\to\infty}\frac{j}{j+k}$ and $\lim_{k\to\infty}\frac{j}{j+k}$ exist for each fixed k and j, respectively. This cautions about trying to use these limits to infer convergence of the double sequence.

3. The double sequence $(\frac{(-1)^j}{k})_{j,k\in\mathbb{Z}_{>0}}$ is easily seen to converge to 0. However, the limit $\lim_{j\to\infty}\frac{(-1)^j}{k}$ does not exist for any fixed k. Therefore, one needs condition (ii) in Proposition 2.3.20 and conditions (ii) and (iii) in Proposition 2.3.21 in order for the results to be valid. •

2.3.6 Algebraic operations on sequences

It is of frequent interest to add, multiply, or divide sequences and series. In such cases, one would like to ensure that convergence of the sequences or series is sufficient to ensure convergence of the sum, product, or quotient. In this section we address this matter.

2.3.23 **Proposition (Algebraic operations on sequences)** *Let* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *and* $(y_j)_{j\in\mathbb{Z}_{>0}}$ *be sequences converging to* s_0 *and* t_0*, respectively, and let* $\alpha \in \mathbb{R}$*. Then the following statements hold:*

(i) the sequence $(\alpha x_j)_{j\in\mathbb{Z}_{>0}}$ *converges to* αs_0*;*

(ii) the sequence $(x_j + y_j)_{j\in\mathbb{Z}_{>0}}$ *converges to* $s_0 + t_0$*;*

(iii) the sequence $(x_j y_j)_{j\in\mathbb{Z}_{>0}}$ *converges to* $s_0 t_0$*;*

(iv) if, for all $j \in \mathbb{Z}_{>0}$, $y_j \neq 0$ *and if* $s_0 \neq 0$*, then the sequence* $(\frac{x_j}{y_j})_{j\in\mathbb{Z}_{>0}}$ *converges to* $\frac{s_0}{t_0}$*.*

Proof (i) The result is trivially true for $a = 0$, so let us suppose that $a \neq 0$. Let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $|x_j - s_0| < \frac{\epsilon}{|\alpha|}$. Then, for $j \geq N$,

$$|\alpha x_j - \alpha s_0| = |\alpha||x_j - s_0| < \epsilon.$$

(ii) Let $\epsilon \in \mathbb{R}_{>0}$ and take $N_1, N_2 \in \mathbb{Z}_{>0}$ such that

$$|x_j - s_0| < \tfrac{\epsilon}{2}, \quad j \geq N_1, \qquad |y_j - t_0| < \tfrac{\epsilon}{2}, \quad j \geq N_2.$$

Then, for $j \geq \max\{N_1, N_2\}$,

$$|x_j + y_j - (s_0 + t_0)| \leq |x_j - s_0| + |y_j - t_0| = \epsilon,$$

using the triangle inequality.

(iii) Let $\epsilon \in \mathbb{R}_{>0}$ and define $N_1, N_2, N_3 \in \mathbb{Z}_{>0}$ such that

$$\begin{aligned}
&|x_j - s_0| < 1, \qquad j \geq N_1, \quad\Longrightarrow\quad |x_j| < |s_0| + 1, \qquad j \geq N_1,\\
&|x_j - s_0| < \frac{\epsilon}{2(|t_0| + 1)}, \qquad j \geq N_2,\\
&|y_j - t_0| < \frac{\epsilon}{2(|s_0| + 1)}, \qquad j \geq N_2.
\end{aligned}$$

Then, for $j \geq \max\{N_1, N_2, N_3\}$,

$$\begin{aligned}
|x_j y_j - s_0 t_0| &= |x_j y_j - x_j t_0 + x_j t_0 - s_0 t_0|\\
&= |x_j(y_j - t_0) + t_0(x_j - s_0)|\\
&\leq |x_j||y_j - t_0| + |t_0||x_j - s_0|\\
&\leq (|s_0| + 1)\frac{\epsilon}{2(|s_0| + 1)} + (|t_0| + 1)\frac{\epsilon}{2(|t_0| + 1)} = \epsilon.
\end{aligned}$$

(iv) It suffices using part (iii) to consider the case where $x_j = 1$, $j \in \mathbb{Z}_{>0}$. For $\epsilon \in \mathbb{R}_{>0}$ take $N_1.N_2 \in \mathbb{Z}_{>0}$ such that

$$\begin{aligned}
&|y_j - t_0| < \frac{|t_0|}{2}, \qquad j \geq N_1, \quad\Longrightarrow\quad |y_j| > \frac{|t_0|}{2}, \qquad j \geq N_1,\\
&|y_j - t_0| < \frac{|t_0|^2\epsilon}{2}, \qquad j \geq N_2.
\end{aligned}$$

Then, for $j \geq \max\{N_1, N_2\}$,

$$\left|\frac{1}{y_j} - \frac{1}{t_0}\right| = \left|\frac{y_j - t_0}{y_j t_0}\right| \leq \frac{|t_0|^2\epsilon}{2}\frac{2}{|t_0|}\frac{1}{|t_0|} = \epsilon,$$

as desired. ■

As we saw in the statement of Proposition 2.2.1, the restriction in part (iv) that $y_j \neq 0$ for all $j \in \mathbb{Z}_{>0}$ is not a real restriction. The salient restriction is that the sequence $(y_j)_{j\in\mathbb{Z}_{>0}}$ not converge to 0.

2.3.7 Convergence using ℝ-nets

Up to this point in this section we have talked about convergence of sequences. However, in practice it is often useful to take limits of more general objects where the index set is not $\mathbb{Z}_{>0}$, but a subset of $\mathbb{R}$. In Section 1.4.4 we introduced a generalisation of sequences called nets. In this section we consider particular cases of nets, called $\mathbb{R}$-nets, that arise commonly when dealing with real numbers and subsets of real numbers. These will be particularly useful when considering the relationships between limits and functions. As we shall see, this slightly more general notion of convergence can be reduced to standard convergence of sequences. We comment that the notions of convergence in this section can be generalised to general nets, and we refer the reader to ***missing stuff*** for details.

Our objective is to understand what is meant by an expression like $\lim_{x\to a} \phi(a)$, where $\phi\colon A \to \mathbb{R}$ is a map from a subset A of $\mathbb{R}$ to $\mathbb{R}$. We will mainly be interested in subsets A of a rather specific form. However, we consider the general case so as to cover all situations that might arise.

2.3.24 Definition (ℝ-directed set) A $\mathbb{R}$-*directed set* is a pair $D = (A, \preceq)$ where the partial order $\preceq$ is defined by $x \preceq y$ if either

(i) $x \leq y$,

(ii) $x \geq y$, or

(iii) there exists $x_0 \in \mathbb{R}$ such that $|x - x_0| \leq |y - x_0|$ (we abbreviate this relation as $x \preceq_{x_0} y$). •

Note that if $D = (A, \preceq)$ is a $\mathbb{R}$-directed set, then it is indeed a directed set because, corresponding to the three cases of the definition,

1. if $x, y \in A$, then $z = \max\{x, y\}$ has the property that $x \preceq z$ and $y \preceq z$ (for the first case in the definition),

2. if $x, y \in A$, then $z = \min\{x, y\}$ has the property that $x \preceq z$ and $y \preceq z$ (for the second case in the definition), or

3. if $x, y \in A$ then, taking z to satisfy $|z - x_0| = \min\{|x - x_0|, |y - x_0|\}$, we have $x \preceq z$ and $y \preceq z$ (for the third case of the definition).

Let us give some examples to illustrate the sort of phenomenon one can see for $\mathbb{R}$-directed sets.

2.3.25 Examples ($\mathbb{R}$-directed sets)

1. Let us take the $\mathbb{R}$-directed set $([0,1], \leq)$. Here we see that, for any $x, y \in [0,1]$, we have $x \leq 1$ and $y \leq 1$.
2. Next take the $\mathbb{R}$-directed set $([0,1), \leq)$. Here, there is no element z of $[0,1)$ for which $x \leq z$ and $y \leq z$ for every $x, y \in [0,1)$. However, it obviously holds that $x \leq 1$ and $y \leq 1$ for every $x, y \in [0,1)$.
3. Next we consider the $\mathbb{R}$ directed set $([0,\infty), \geq)$. Here we see that, for any $x, y \in [0,\infty)$, $x \geq 0$ and $y \geq 0$.
4. Next we consider the $\mathbb{R}$ directed set $((0,\infty), \geq)$. Here we see that there is no element $z \in (0,\infty)$ such that, for every $x, y \in (0,\infty)$, $x \geq z$ and $y \geq z$. However, it is true that $x \geq 0$ and $y \geq 0$ for every $x, y \in (0,\infty)$.
5. Now we take the $\mathbb{R}$-directed set $([0,\infty), \leq)$. Here we see that there is no element $z \in [0,\infty)$ such that $x \leq z$ and $y \leq z$ for every $x, y \in [0,\infty)$. Moreover, there is also no element $z \in \mathbb{R}$ for which $x \leq z$ and $y \leq z$ for every $x, y \in [0,\infty)$.
6. Next we take the $\mathbb{R}$-directed set $(\mathbb{Z}, \leq)$. As in the preceding example, there is no element $z \in [0,\infty)$ such that $x \leq z$ and $y \leq z$ for every $x, y \in [0,\infty)$. Moreover, there is also no element $z \in \mathbb{R}$ for which $x \leq z$ and $y \leq z$ for every $x, y \in [0,\infty)$.
7. Now consider the $\mathbb{R}$-directed set $(\mathbb{R}, \leq_0)$. Note that $0 \in \mathbb{R}$ has the property that, for any $x, y \in \mathbb{R}$, $x \leq_0 0$ and $y \leq_0 0$.
8. Similar to the preceding example, consider the $\mathbb{R}$-directed set $(\mathbb{R} \setminus \{0\}, \leq_0)$. Here there is no element $z \in \mathbb{R} \setminus \{0\}$ such that $x \leq_0 z$ and $y \leq_0 z$ for every $x, y \in \mathbb{R} \setminus \{0\}$. However, we clearly have $x \leq_0 0$ and $y \leq_0 0$ for every $x, y \in \mathbb{R} \setminus \{0\}$. •

The examples may seem a little silly, but this is just because the notion of a $\mathbb{R}$-directed set is, in and of itself, not so interesting. What is more interesting is the following notion.

2.3.26 Definition ($\mathbb{R}$-net, convergence in $\mathbb{R}$-nets) If $D = (A, \preceq)$ is a $\mathbb{R}$-directed set, a $\mathbb{R}$-***net*** in D is a map $\phi\colon A \to \mathbb{R}$. A $\mathbb{R}$-net $\phi\colon A \to \mathbb{R}$ in a $\mathbb{R}$-directed set $D = (A, \preceq)$

(i) ***converges*** to $s_0 \in \mathbb{R}$ if, for any $\epsilon \in \mathbb{R}_{>0}$, there exists $x \in A$ such that $|\phi(y) - s_0| < \epsilon$ for any $y \in A$ satisfying $x \preceq y$,

(ii) has s_0 as a ***limit*** if it converges to s_0, and we write $s_0 = \lim_D \phi$,

(iii) ***diverges*** if it does not converge,

(iv) ***diverges to ∞*** ((resp. ***diverges to −∞***, and we write $\lim_D \phi = \infty$ (resp. $\lim_D \phi = -\infty$), if, for each $M \in \mathbb{R}_{>0}$, there exists $x \in A$ such that $\phi(y) > M$ (resp. $\phi(y) < -M$) for every $y \in A$ for which $x \preceq y$,

(v) has a limit that ***exists*** if $\lim_D \phi \in \mathbb{R}$, and

(vi) is ***oscillatory*** if the limit of the $\mathbb{R}$-net does not exist, does not diverge to ∞, and does not diverge to $-\infty$. •

2.3.27 Notation (Limits of $\mathbb{R}$-nets) The importance $\mathbb{R}$-nets can now be illustrated by showing how they give rise to a collection of convergence phenomenon. Let us look at various cases for convergence of a $\mathbb{R}$-net in a $\mathbb{R}$-directed set $D = (A, \preceq)$.

(i) $\preceq = \leq$: Here there are two subcases to consider.

(a) $\sup A = x_0 < \infty$: In this case we write $\lim_D \phi = \lim_{x \uparrow x_0} \phi(x)$.

(b) $\sup A = \infty$: In this case we write $\lim_D \phi = \lim_{x \to \infty} \phi(x)$.

(ii) $\preceq = \geq$: Again we have two subcases.

(a) $\inf A = x_0 > -\infty$: In this case we write $\lim_D \phi = \lim_{x \downarrow x_0} \phi(x)$.

(b) $\inf A = -\infty$: In this case we write $\lim_D \phi = \lim_{x \to -\infty} \phi(x)$.

(iii) $\preceq = \leq_{x_0}$: There are three subcases here that we wish to distinguish.

(a) $\sup A = x_0$: Here we denote $\lim_D \phi = \lim_{x \uparrow x_0} \phi(x)$.

(b) $\inf A = x_0$: Here we denote $\lim_D \phi = \lim_{x \downarrow x_0} \phi(x)$.

(c) $x_0 \notin \{\inf A, \sup A\}$: Here we denote $\lim_D \phi = \lim_{x \to x_0} \phi(x)$. •

In the case when the directed set is an interval, we have the following notation that unifies the various limit notations for this special often encountered case.

2.3.28 Notation (Limit in an interval) Let $I \subseteq \mathbb{R}$ be an interval, let $\phi \colon I \to \mathbb{R}$ be a map, and let $a \in I$. We define $\lim_{x \to_I a} \phi(x)$ by

(i) $\lim_{x \to_I a} \phi(x) = \lim_{x \uparrow a} \phi(x)$ if $a = \sup I$,

(ii) $\lim_{x \to_I a} \phi(x) = \lim_{x \downarrow a} \phi(x)$ if $a = \inf I$, and

(iii) $\lim_{x \to_I a} \phi(x) = \lim_{x \to a} \phi(x)$ otherwise. •

We expect that most readers will be familiar with the idea here, even if the notation is not conventional. Let us also give the notation a precise characterisation in terms of limits of sequences in the case when the point x_0 is in the closure of the set A.

2.3.29 Proposition (Convergence in $\mathbb{R}$-nets in terms of sequences) *Let* $(\mathrm{A}, \preceq)$ *be a* $\mathbb{R}$-*directed set and let* $\phi \colon \mathrm{A} \to \mathbb{R}$ *be a* $\mathbb{R}$-*net in* $(\mathrm{A}, \preceq)$. *Then, corresponding to the cases and subcases of Notation 2.3.27, we have the following statements:*

(i) *(a)* *if* $\mathrm{x}_0 \in \mathrm{cl}(\mathrm{A})$, *the following statements are equivalent:*

I. $\lim_{\mathrm{x} \uparrow \mathrm{x}_0} \phi(\mathrm{x}) = \mathrm{s}_0$;

II. $\lim_{\mathrm{j} \to \infty} \phi(\mathrm{x}_\mathrm{j}) = \mathrm{s}_0$ *for every sequence* $(\mathrm{x}_\mathrm{j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{\mathrm{j} \to \infty} \mathrm{x}_\mathrm{j} = \mathrm{x}_0$;

(b) *the following statements are equivalent:*

I. $\lim_{\mathrm{x} \to \infty} \phi(\mathrm{x}) = \mathrm{s}_0$;

II. $\lim_{\mathrm{j} \to \infty} \phi(\mathrm{x}_\mathrm{j}) = \mathrm{s}_0$ *for every sequence* $(\mathrm{x}_\mathrm{j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{\mathrm{j} \to \infty} \mathrm{x}_\mathrm{j} = \infty$;

(ii) *(a)* *if* $\mathrm{x}_0 \in \mathrm{cl}(\mathrm{A})$, *the following statements are equivalent:*

I. $\lim_{\mathrm{x} \downarrow \mathrm{x}_0} \phi(\mathrm{x}) = \mathrm{s}_0$;

II. $\lim_{j\to\infty}\phi(x_j) = s_0$ *for every sequence* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{j\to\infty} x_j = x_0$*;*

(b) *the following statements are equivalent:*

I. $\lim_{x\to-\infty}\phi(x) = s_0$*;*

II. $\lim_{j\to\infty}\phi(x_j) = s_0$ *for every sequence* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{j\to\infty} x_j = -\infty$*;*

(iii) (a) *if* $x_0 \in \mathrm{cl}(A)$*, the following statements are equivalent:*

I. $\lim_{x\uparrow x_0}\phi(x) = s_0$*;*

II. $\lim_{j\to\infty}\phi(x_j) = s_0$ *for every sequence* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{j\to\infty} x_j = x_0$*;*

(b) *if* $x_0 \in \mathrm{cl}(A)$*, the following statements are equivalent:*

I. $\lim_{x\downarrow x_0}\phi(x) = s_0$*;*

II. $\lim_{j\to\infty}\phi(x_j) = s_0$ *for every sequence* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{j\to\infty} x_j = x_0$*;*

(c) *the following statements are equivalent:*

I. $\lim_{x\to\infty}\phi(x) = s_0$*;*

II. $\lim_{j\to\infty}\phi(x_j) = s_0$ *for every sequence* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *in* A *satisfying* $\lim_{j\to\infty} x_j = \infty$*;*

Proof These statements are all proved in essentially the same way, so let us prove just, say, part (i a).

First suppose that $\lim_{x\uparrow x_0}\phi(x) = s_0$, and let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in A converging to x_0. Let $\epsilon \in \mathbb{R}_{>0}$ and choose $x \in A$ such that $|\phi(y) - s_0| < \epsilon$ whenever $y \in A$ satisfies $x \leq y$. Then, since $\lim_{j\to\infty} x_j = x_0$, there exists $N \in \mathbb{Z}_{>0}$ such that $x \leq x_j$ for all $j \geq N$. Clearly, $|\phi(x_j) - s_0| < \epsilon$, so giving convergence of $(\phi(x_j))_{j\in\mathbb{Z}_{>0}}$ to s_0 for every sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ in A converging to x_0.

For the converse, suppose that $\lim_{x\uparrow x_0}\phi(x) \neq s_0$. Then there exists $\epsilon \in \mathbb{R}_{>0}$ such that, for any $x \in A$, we have a $y \in A$ with $x \leq y$ for which $|\phi(y) - s_0| \geq \epsilon$. Since $x_0 \in \mathrm{cl}(A)$ it follows that, for any $j \in \mathbb{Z}_{>0}$, there exists $x_j \in \mathsf{B}(\frac{1}{j}, x_0) \cap A$ such that $|\phi(x_j) - s_0| \geq \epsilon$. Thus the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ in A converging to x_0 has the property that $(\phi(x_j))_{j\in\mathbb{Z}_{>0}}$ does not converge to s_0. ■

Of course, similar conclusions hold when "convergence to s_0" is replaced with "divergence," "convergence to ∞," "convergence to $-\infty$," or "oscillatory." We leave the precise statements to the reader.

Let us give some examples to illustrate that this is all really nothing new.

2.3.30 Examples (Convergence in $\mathbb{R}$-nets)

1. Consider the $\mathbb{R}$-directed set $([0,\infty), \leq)$ and the corresponding $\mathbb{R}$-net ϕ defined by $\phi(x) = \frac{1}{1+x^2}$. This $\mathbb{R}$-net then converges to 0. Let us verify this using the formal definition of convergence of a $\mathbb{R}$-net. For $\epsilon \in \mathbb{R}_{>0}$ choose $x > 0$ such that $x^2 = \frac{1}{\epsilon} > \frac{1}{\epsilon} - 1$. Then, if $x \leq y$, we have

$$\left|\frac{1}{1+y^2} - 0\right| < \frac{1}{1+x^2} < \epsilon,$$

giving convergence to $\lim_{x\to\infty}\phi(x) = 0$ as stated.

2. Next consider the $\mathbb{R}$-directed set $((0,1],\geq)$ and the corresponding $\mathbb{R}$-net ϕ defined by $\phi(x) = x\sin\frac{1}{x}$. We claim that this $\mathbb{R}$-net converges to 0. To see this, let $\epsilon\in\mathbb{R}_{>0}$ and let $x\in(0,\epsilon)$. Then we have, for $x\geq y$,
$$\left|y\sin\tfrac{1}{y} - 0\right| = y \leq x < \epsilon,$$
giving $\lim_{x\downarrow 0}\phi(x) = 0$ as desired.
3. Consider the $\mathbb{R}$-directed set $([0,\infty),\leq)$ and the associated $\mathbb{R}$-net ϕ defined by $\phi(x) = x$. In this case we have $\lim_{x\to\infty}\phi(x) = \infty$.
4. Consider the $\mathbb{R}$-directed set $([0,\infty),\leq)$ and the associated $\mathbb{R}$-net ϕ defined by $\phi(x) = x\sin x$. In this case, due to the oscillatory nature of sin, $\lim_{x\to\infty}\phi(x)$ does not exist, nor does it diverge to either ∞ or $-\infty$.
5. Take the $\mathbb{R}$-directed set $(\mathbb{R}\setminus\{0\},\leq_0)$. Define the $\mathbb{R}$-net ϕ by $\phi(x) = x$. Clearly, $\lim_{x\to 0}\phi(x) = 0$. •

There are also generalisations of lim sup and lim inf to $\mathbb{R}$-nets. We let $D = (A,\preceq)$ be a $\mathbb{R}$-directed set and let $\phi\colon A\to\mathbb{R}$ be a $\mathbb{R}$-net in this $\mathbb{R}$-directed set. We denote by $\sup_D\phi, \inf_D\phi\colon A\to\mathbb{R}$ the $\mathbb{R}$-nets in D given by
$$\sup_D\phi(x) = \sup\{\phi(y) \mid x\preceq y\},\quad \inf_D\phi(x) = \inf\{\phi(y) \mid x\preceq y\}.$$
Then we define
$$\limsup_D\phi = \lim_D\sup_D\phi,\quad \liminf_D\phi = \lim_D\inf_D\phi.$$
These allow us to talk of limits in cases where limits in the usual sense to not exist. Let us consider this via an example.

2.3.31 Example (lim sup and lim inf in $\mathbb{R}$-nets) We consider the $\mathbb{R}$-directed set $D = ([0,\infty),\leq)$ and let ϕ be the $\mathbb{R}$-net defined by $\phi(x) = \mathrm{e}^{-x} + \sin x$.[5] We claim that $\limsup_D\phi = 1$ and that $\liminf_D\phi = -1$. Let us prove the first claim, and leave the second as an exercise. We then have
$$\sup_D\phi(x) = \sup\{\mathrm{e}^{-y} + \sin y \mid x\leq y\} = \mathrm{e}^{-x} + 1.$$
First note that $\sup_D\phi(x)\geq 1$ for every $x\in[0,\infty)$, and so $\limsup_D\phi\geq 1$. Now let $\epsilon\in\mathbb{R}_{>0}$ and take $x > \log\epsilon$. Then, for any $y\geq x$,
$$\sup_D\phi(y) = \mathrm{e}^{-y} + 1 \leq 1+\epsilon.$$
Therefore, $\limsup_D\phi\leq 1$, and so $\limsup_D\phi = 1$, as desired. •

[5] We have not yet defined e^{-x} or $\sin x$. The reader who is unable to go on without knowing what these functions really are can skip ahead to Section 3.6.

2.3.8 A first glimpse of Landau symbols

In this section we introduce for the first time the so-called Landau symbols. These provide commonly used notation for when two functions behave "asymptotically" the same. Given our development of $\mathbb{R}$-nets in the preceding section, it is easy for us to be fairly precise here. We also warn the reader that the Landau symbols often get used in an imprecise or vague way. We shall try to avoid such usage.

We begin with the definition.

2.3.32 Definition (Landau symbols "O" and "o") Let $D = (A, \preceq)$ be a $\mathbb{R}$-directed set and let $\phi\colon A \to \mathbb{R}$.

(i) Denote by $O_D(\phi)$ the functions $\psi\colon A \to \mathbb{R}$ for which there exists $x_0 \in A$ and $M \in \mathbb{R}_{>0}$ such that $|\psi(x)| \le M|\phi(x)|$ for $x \in A$ satisfying $x_0 \preceq x$.

(ii) Denote by $o_D(\phi)$ the functions $\psi\colon A \to \mathbb{R}$ such that, for any $\epsilon \in \mathbb{R}_{>0}$, there exists $x_0 \in A$ such that $|\psi(x)| < \epsilon|\phi(x)|$ for $x \in A$ satisfying $x_0 \preceq x$.

If $\psi \in O_D(\phi)$ (resp. $\psi \in o_D(\phi)$) then we say that ψ is ***big oh of ϕ*** (resp. ***little oh of ϕ***). •

It is very common to see simply $O(\phi)$ and $o(\phi)$ in place of $O_D(\phi)$ and $o_D(\phi)$. This is because the most common situation for using this notation is in the case when $\sup A = \infty$ and $\preceq=\le$. In such cases, the notation indicates means, essentially, that $\psi \in O(\phi)$ if ψ has "size" no larger than ϕ for large values of the argument and that $\psi \in o(\phi)$ if ψ is "small" compared to ϕ for large values of the argument. However, we shall use the Landau symbols in other cases, so we allow the possibility of explicitly including the $\mathbb{R}$-directed set in our notation for the sake of clarity.

It is often the case that the comparison function ϕ is positive on A. In such cases, one can give a somewhat more concrete characterisation of O_D and o_D.

2.3.33 Proposition (Alternative characterisation of Landau symbols) *Let* $\mathrm{D} = (\mathrm{A}, \preceq)$ *be a* $\mathbb{R}$*-directed set, and let* $\phi\colon \mathrm{A} \to \mathbb{R}_{>0}$ *and* $\psi\colon \mathrm{A} \to \mathbb{R}$*. Then*

(i) $\psi \in \mathrm{O}_\mathrm{D}(\phi)$ *if and only if* $\limsup_\mathrm{D} \frac{\psi}{\phi} < \infty$ *and*

(ii) $\psi \in \mathrm{o}_\mathrm{D}(\phi)$ *if and only if* $\lim_\mathrm{D} \frac{\psi}{\phi} = 0$*.*

Proof We leave this as Exercise 2.3.6. ■

Let us give some common examples of where the Landau symbols are used. Some examples will make use of ideas we have not yet discussed, but which we imagine are familiar to most readers.

2.3.34 Examples (Landau symbols)

1. Let $I \subseteq \mathbb{R}$ be an interval for which $x_0 \in I$ and let $f\colon I \to \mathbb{R}$. Consider the $\mathbb{R}$-directed set $D = (I \setminus \{x_0\}, \preceq_{x_0})$ and the $\mathbb{R}$-net ϕ in D given by $\phi(x) = 1$. Define $g_{f,x_0}\colon I \to \mathbb{R}$ by $g_{f,x_0}(x) = f(x_0)$. We claim that f is continuous at x_0 if and only if

$f - g_{f,x_0} \in o_D(\phi)$. Indeed, by Theorem 3.1.3 we have that f is continuous at x_0 if and only if

$$\begin{aligned} & \lim_{x \to_I x_0} f(x) = f(x_0) \\ \implies \quad & \lim_{x \to_I x_0} (f(x) - g_{f,x_0}(x)) = 0 \\ \implies \quad & \lim_{x \to_I x_0} \frac{(f(x) - g_{f,x_0}(x))}{\phi(x)} = 0 \\ \implies \quad & f - g_{f,x_0} \in o_D(\phi). \end{aligned}$$

The idea is that f is continuous at x_0 if and only if f is "approximately constant" near x_0.

2. Let $I \subseteq \mathbb{R}$ be an interval for which $x_0 \in I$ and let $f\colon I \to \mathbb{R}$. For $L \in \mathbb{R}$ define $g_{f,x_0,L}\colon I \setminus \{x_0\} \to \mathbb{R}$ by
$$g_{x_0,L}(x) = f(x_0) + L(x - x_0).$$
Consider the $\mathbb{R}$-directed set $D = (I \setminus \{x_0\}, \leq_{x_0})$, and define $\phi\colon I \setminus \{x_0\} \to \mathbb{R}_{>0}$ by $\phi(x) = |x - x_0|$. Then we claim that f is differentiable at x_0 with derivative $f'(x_0) = L$ if and only if $f - g_{f,x_0,L} \in o_D(\phi)$. Indeed, by definition, f is differentiable at x_0 with derivative $f'(x_0) = L$ if and only if, then
$$\begin{aligned} & \lim_{x \to_I x_0} \frac{f(x) - f(x_0)}{x - x_0} = L \\ \iff \quad & \lim_{x \to_I x_0} \frac{1}{x - x_0}\big(f(x) - g_{f,x_0,L}(x)\big) = 0 \\ \iff \quad & \lim_{x \to_I x_0} \frac{1}{|x - x_0|}\big(f(x) - g_{f,x_0,L}(x)\big) = 0 \\ \iff \quad & f(x) - g_{f,x_0,L}(x) \in o_D(\phi), \end{aligned}$$
using Proposition 2.3.33. The idea is that f is differentiable at x_0 if and only if f is "nearly linear" at x_0.

3. We can generalise the preceding two examples. Let $I \subseteq \mathbb{R}$ be an interval, let $x_0 \in I$, and consider the $\mathbb{R}$-directed set $(I \setminus \{x_0\}, \leq_{x_0})$. For $m \in \mathbb{Z}_{\geq 0}$ define the $\mathbb{R}$-net ϕ_m in D by $\phi_m(x) = |x - x_0|^m$. We shall say that a function $f\colon I \to \mathbb{R}$ ***vanishes to order*** $\mathbf{m}$ ***at*** $\mathbf{x_0}$ if $f \in O_D(\phi_m)$. Moreover, f is m-times differentiable at x_0 with $f^{(j)}(x_0) alpha_j$, $j \in \{0, 1, \dots, m\}$, if and only if $f - g_{f,x_0,\boldsymbol{\alpha}} \in o_D(\phi_m)$, where
$$g_{f,x_0,\boldsymbol{\alpha}}(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_m x^m.$$

4. One of the common places where Landau symbols are used is in the analysis of the complexity of algorithms. An algorithm, loosely speaking, takes some input data, performs operations on the data, and gives an outcome. A very simple example of an algorithm is the multiplication of two square matrices, and we will use this simple example to illustrate our discussion. It is assumed that the size of the input data is measured by an integer N. For example, for

the multiplication of square matrices, this integer is the size of the matrices. The complexity of an algorithm is then determined by the number of steps, denoted by, say, $\psi(N)$, of a certain type in the algorithm. For example, for the multiplication of square matrices, this number is normally taken to be the number of multiplications that are needed, and this is easily seen to be no more than N^2. To describe the complexity of the algorithm, one finds uses Landau symbols in the following way. First of all, we use the $\mathbb{R}$-directed set $D = (\mathbb{Z}_{>0}, \leq)$. If $\phi\colon \mathbb{Z}_{>0} \to \mathbb{R}_{>0}$ is such that $\psi \in O_D(\phi)$, then we say the algorithm ***is* $\mathbf{O(\phi)}$**. For example, matrix multiplication is $O(N^2)$.

In Theorem 14.2.20 we show that the computational complexity of the so-called Cooley–Tukey algorithm for computing the FFT is $O(N \log N)$.

Since we are talking about computational complexity of algorithms, it is a good time to make mention of an important problem in the theory of computational complexity. This discussion is limited to so-called decision algorithms, where the outcome is an affirmative or negative declaration about some problem, e.g., is the determinant of a matrix bounded by some number. For such an algorithm, a ***verification algorithm*** is an algorithm that checks whether given input data does indeed give an affirmative answer. Denote by ***P*** the class of algorithms that are $O(N^m)$ for some $m \in \mathbb{Z}_{>0}$. Such algorithms are known as ***polynomial time*** algorithms. Denote by ***NP*** the class of algorithms for which there exists a verification algorithm that is $O(N^m)$ for some $m \in \mathbb{Z}_{>0}$. An important unresolved question is, "Does P=NP?" •

2.3.9 Notes

Citation for Dedekind cuts.

Exercises

2.3.1 Show that if $(x_j)_{j\in\mathbb{Z}_{>0}}$ is a sequence in $\mathbb{R}$ and if $\lim_{j\to\infty} x_j = x_0$ and $\lim_{j\to\infty} x_j = x_0'$, then $x_0 = x_0'$.

2.3.2 Answer the following questions:

(a) find a subset $S \subseteq \mathbb{Q}$ that possesses an upper bound in $\mathbb{Q}$, but which has no least element;

(b) find a bounded monotonic sequence in $\mathbb{Q}$ that does not converge in $\mathbb{Q}$.

2.3.3 Do the following.

(a) Find a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which converges in $\mathbb{R}$.

(b) Find a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which diverges to ∞.

(c) Find a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which diverges to $-\infty$.

(d) Find a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which is oscillatory.

2.3.4 *missing stuff*

In the next exercise you will show that the property that a bounded, monotonically increasing sequence converges implies that Cauchy sequences converge. This completes the argument needed to prove the theorem stated in Aside 2.3.9 concerning characterisations of complete ordered fields.

2.3.5 Assume that every bounded, monotonically increasing sequence in $\mathbb{R}$ converges, and using this show that every Cauchy sequence in $\mathbb{R}$ converges using an argument as follows.

1. Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence.
2. Let $I_0 = [a, b]$ be an interval that contains all elements of $(x_j)_{j\in\mathbb{Z}_{>0}}$ (why is this possible?)
3. Split $[a, b]$ into two equal length closed intervals, and argue that in at least one of these there is an infinite number of points from the sequence. Call this interval I_1 and let $x_{k_i} \in (x_j)_{j\in\mathbb{Z}_{>0}} \cap I_1$.
4. Repeat the process for I_1 to find an interval I_2 which contains an infinite number of points from the sequence. Let $x_{k_2} \in (x_j)_{j\in\mathbb{Z}_{>0}} \cap I_2$.
5. Carry on doing this to arrive at a sequence $(x_{k_j})_{j\in\mathbb{Z}_{>0}}$ of points in $\mathbb{R}$ and a sequence $(I_j)_{j\in\mathbb{Z}_{>0}}$.
6. Argue that the sequence of left endpoints of the intervals $(I_j)_{j\in\mathbb{Z}_{>0}}$ is a bounded monotonically increasing sequence, and that the sequence of right endpoints is a bounded monotonically decreasing sequence. and so both converge.
7. Show that they converge to the same number, and that the sequence $(x_{k_j})_{j\in\mathbb{Z}_{>0}}$ also converges to this limit.
8. Show that the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ converges to this limit.

2.3.6 Prove Proposition 2.3.33.

Section 2.4

Series in $\mathbb{R}$

From a sequence $(x_j)_{j\in\mathbb{R}}$ in $\mathbb{R}$, one can consider, in principle, the infinite sum $\sum_{j=1}^{\infty} x_j$. Of course, such a sum *a priori* makes no sense. However, as we shall see in Chapter 8, such infinite sums are important for characterising certain discrete-time signal spaces. Moreover, such sums come up frequently in many places in analysis. In this section we outline some of the principle properties of these sums.

Do I need to read this section? Most readers will probably have seen much of the material in this section in their introductory calculus course. What might be new for some readers is the fairly careful discussion in Theorem 2.4.5 of the difference between convergence and absolute convergence of series. Since absolute convergence will be of importance to us, it might be worth understanding in what ways it is different from convergence. The material in Section 2.4.7 can be regarded as optional until it is needed during the course of reading other material in the text. •

2.4.1 Definitions and properties of series

A ***series*** in $\mathbb{R}$ is an expression of the form

$$S = \sum_{j=1}^{\infty} x_j, \tag{2.5}$$

where $x_j \in \mathbb{R}$, $j \in \mathbb{Z}_{>0}$. Of course, the problem with this "definition" is that the expression (2.5) is meaningless as an element of $\mathbb{R}$ unless it possesses additional features. For example, if $x_j = 1$, $j \in \mathbb{Z}_{>0}$, then the sum is infinite. Also, if $x_j = (-1)^j$, $j \in \mathbb{Z}_{>0}$, then it is not clear what the sum is: perhaps it is 0 or perhaps it is 1. Therefore, to be precise, a series is prescribed by the sequence of numbers $(x_j)_{j\in\mathbb{Z}_{>0}}$, and is represented in the form (2.5) in order to distinguish it from the sequence with the same terms.

If the expression (2.5) is to have meaning as a number, we need some sort of condition placed on the terms in the series.

2.4.1 Definition (Convergence and absolute convergence of series) Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$ and consider the series

$$S = \sum_{j=1}^{\infty} x_j.$$

The corresponding sequence of ***partial sums*** is the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ defined by

$$S_k = \sum_{j=1}^{k} x_j.$$

Let $s_0 \in \mathbb{R}$. The series:

(i) ***converges to*** $\mathbf{s_0}$, and we write $\sum_{j=1}^{\infty} x_j = s_0$, if the sequence of partial sums converges to s_0;

(ii) has s_0 as a ***limit*** if it converges to s_0;

(iii) is ***convergent*** if it converges to some member of $\mathbb{R}$;

(iv) ***converges absolutely***, or is ***absolutely convergent***, if the series

$$\sum_{j=1}^{\infty} |x_j|$$

converges;

(v) ***converges conditionally***, or is ***conditionally convergent***, if it is convergent, but not absolutely convergent;

(vi) ***diverges*** if it does not converge;

(vii) ***diverges to*** $\boldsymbol{\infty}$ (resp. ***diverges to*** $\boldsymbol{-\infty}$), and we write $\sum_{j=1}^{\infty} x_j = \infty$ (resp. $\sum_{j=1}^{\infty} x_j = -\infty$), if the sequence of partial sums diverges to ∞ (resp. diverges to $-\infty$);

(viii) has a limit that ***exists*** if $\lim_{j\to\infty} S_j \in \mathbb{R}$;

(ix) is ***oscillatory*** if the sequence of partial sums is oscillatory. •

Let us consider some examples of series in $\mathbb{R}$.

2.4.2 Examples (Series in $\mathbb{R}$)

1. First we consider the ***geometric series*** $\sum_{j=1}^{\infty} x^{j-1}$ for $x \in \mathbb{R}$. We claim that this series converges if and only if $|x| < 1$. To prove this we claim that the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ of partial sums is defined by

$$S_k = \begin{cases} \frac{1-x^{k+1}}{1-x}, & x \neq 1, \\ k, & x = 1. \end{cases}$$

The conclusion is obvious for $x = 1$, so we can suppose that $x \neq 1$. The conclusion is obvious for $k = 1$, so suppose it true for $j \in \{1, \dots, k\}$. Then

$$S_{k+1} = \sum_{j=1}^{k+1} x^j = x^{k+1} + \frac{1-x^{k+1}}{1-x} = \frac{x^{k+1} - x^{k+2} + 1 - x^{k+1}}{1-x} = \frac{1-x^{k+2}}{1-x},$$

as desired. It is clear, then, that if $x = 1$ then the series diverges to ∞. If $x = -1$ then the series is directly checked to be oscillatory; the sequence of partial sums is $\{1, 0, 1, \dots\}$. For $x > 1$ we have

$$\lim_{k\to\infty} S_k = \lim_{k\to\infty} \frac{1-x^{k+1}}{1-x} = \infty,$$

showing that the series diverges to ∞ in this case. For $x < -1$ it is easy to see that the sequence of partial sums is oscillatory, but increasing in magnitude.

This leaves the case when $|x| < 1$. Here, since the sequence $(x^{k+1})_{k\in\mathbb{Z}_{>0}}$ converges to zero, the sequence of partial sums also converges, and converges to $\frac{1}{1-x}$. (We have used the results concerning the swapping of limits with algebraic operations as described in Section 2.3.6.)

2. We claim that the series $\sum_{j=1}^{\infty} \frac{1}{j}$ diverges to ∞. To show this, we show that the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ is not upper bounded. To show this, we shall show that $S_{2^k} \geq 1 + \frac{1}{2}k$ for all $k \in \mathbb{Z}_{>0}$. This is true directly when $k = 1$. Next suppose that $S_{2^j} \geq 1 + \frac{1}{2}j$ for $j \in \{1, \dots, k\}$. Then

$$\begin{aligned} S_{2^{k+1}} &= S_{2^k} + \frac{1}{2^k+1} + \frac{1}{2^k+2} + \dots + \frac{1}{2^{k+1}} \\ &\geq 1 + \frac{1}{2}k + \underbrace{\frac{1}{2^{k+1}} + \dots + \frac{1}{2^{k+1}}}_{2^k \text{ terms}} \\ &= 1 + \frac{1}{2}k + \frac{2^k}{2^{k+1}} = 1 + \frac{1}{2}(k+1). \end{aligned}$$

Thus the sequence of partial sums is indeed unbounded, and since it is monotonically increasing, it diverges to ∞, as we first claimed.

3. We claim that the series $S = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}$ converges. To see this, we claim that, for any $m \in \mathbb{Z}_{>0}$, we have

$$S_2 \leq S_4 \leq \dots \leq S_{2m} \leq S_{2m-1} \leq \dots \leq S_3 \leq S_1.$$

That $S_2 \leq S_4 \leq \dots \leq S_{2m}$ follows since $S_{2k} - S_{2k-2} = \frac{1}{2k-1} - \frac{1}{2k} > 0$ for $k \in \mathbb{Z}_{>0}$. That $S_{2m} \leq S_{2m-1}$ follows since $S_{2m-1} - S_{2m} = \frac{1}{2m}$. Finally, $S_{2m-1} \leq \dots \leq S_3 \leq S_1$ since $S_{2k-1} - S_{2k+1} = \frac{1}{2k} - \frac{1}{2k+1} > 0$ for $k \in \mathbb{Z}_{>0}$. Thus the sequences $(S_{2k})_{k\in\mathbb{Z}_{>0}}$ and $(S_{2k-1})_{k\in\mathbb{Z}_{>0}}$ are monotonically increasing and monotonically decreasing, respectively, and their tails are getting closer and closer together since $\lim_{m\to\infty} S_{2m-1} - S_{2m} = \frac{1}{2m} = 0$. By Lemma 2 from the proof of Theorem 2.3.7, it follows that the sequences $(S_{2k})_{k\in\mathbb{Z}_{>0}}$ and $(S_{2k-1})_{k\in\mathbb{Z}_{>0}}$ converge and converge to the same limit. Therefore, the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ converges as well to the same limit. One can moreover show that the limit of the series is $\log 2$, where log denotes the natural logarithm.

Note that we have now shown that the series $\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}$ converges, but does not converge absolutely; therefore, it is conditionally convergent.

4. We next consider the ***harmonic series*** $\sum_{j=1}^{\infty} j^{-k}$ for $k \in \mathbb{Z}_{\geq 0}$. For $k = 1$ this agrees with our example of part 2. We claim that this series converges if and only if $k > 1$. We have already considered the case of $k = 1$. For $k < 1$ we have $j^{-k} \geq j^{-1}$ for $j \in \mathbb{Z}_{>0}$. Therefore,

$$\sum_{j=1}^{\infty} j^{-k} \geq \sum_{j=1}^{\infty} j^{-1} = \infty,$$

showing that the series diverges to ∞.

For $k > 1$ we note that the sequence of partial sums is monotonically increasing. Thus, so show convergence of the series it suffices by Theorem 2.3.8 to show that the sequence of partial sums is bounded above. Let $N \in \mathbb{Z}_{>0}$ and take $j \in \mathbb{Z}_{>0}$ such that $N < 2^j - 1$. Then the Nth partial sum satisfies

$$\begin{aligned} S_N \leq S_{2^j-1} &= 1 + \frac{1}{2^k} + \frac{1}{3^k} + \cdots + \frac{1}{(2^j-1)^k} \\ &= 1 + \underbrace{\Big(\frac{1}{2^k} + \frac{1}{3^k}\Big)}_{\text{2 terms}} + \underbrace{\Big(\frac{1}{4^k} + \cdots + \frac{1}{7^k}\Big)}_{\text{4 terms}} + \cdots + \underbrace{\Big(\frac{1}{(2^{j-1})^k} + \cdots + \frac{1}{(2^j-1)^k}\Big)}_{2^{j-1}\text{ terms}} \\ &< 1 + \frac{2}{2^k} + \frac{4}{4^k} + \cdots + \frac{2^{j-1}}{(2^{j-1})^k} \\ &= 1 + \frac{1}{2^{k-1}} + \Big(\frac{1}{2^{k-1}}\Big)^2 + \cdots + \Big(\frac{1}{2^{k-1}}\Big)^{j-1}. \end{aligned}$$

Now we note that the last expression on the right-hand side is bounded above by the sum $\sum_{j=1}^{\infty}(2^{k-1})^{j-1}$, which is a convergent geometric series as we saw in part 1. This shows that S_N is bounded above by this sum for all N, so showing that the harmonic series converges for $k > 1$.

5. The series $\sum_{j=1}^{\infty}(-1)^{j+1}$ does not converge, and also does not diverge to ∞ or $-\infty$. Therefore, it is oscillatory. •

Let us next explore relationships between the various notions of convergence. First we relate the notions of convergence and absolute convergence in the only possible way, given that the series $\sum_{j=1} \frac{(-1)^{j+1}}{j}$ has been shown to be convergent, but not absolutely convergent.

2.4.3 Proposition (Absolutely convergent series are convergent) *If a series $\sum_{j=1}^{\infty} x_j$ is absolutely convergent, then it is convergent.*

Proof Denote

$$s_k = \sum_{j=1}^{k} x_j, \quad \sigma_k = \sum_{j=1}^{k} |x_j|,$$

and note that $(\sigma_k)_{k\in\mathbb{Z}_{>0}}$ is a Cauchy sequence since the series $\sum_{j=1}^{\infty} x_j$ is absolutely convergent. Thus let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $|\sigma_k - \sigma_l| < \epsilon$ for $k, l \geq N$. For $m > k$ we then have

$$|s_m - s_k| = \Big|\sum_{j=k+1}^{m} x_j\Big| \leq \sum_{j=k+1}^{m} |x_j| = |\sigma_m - \sigma_k| < \epsilon,$$

where we have used Exercise 2.4.3. Thus, for $m > k \geq N$ we have $|s_m - s_k| < \epsilon$, showing that $(s_k)_{k\in\mathbb{Z}_{>0}}$ is a Cauchy sequence, and so convergent by Theorem 2.3.5. ■

The following result is often useful.

2.4.4 Proposition (Swapping summation and absolute value) *For a sequence* $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$*, if the series* $\mathrm{S} = \sum_{j=1}^{\infty} \mathrm{x}_j$ *is absolutely convergent, then*

$$\Big|\sum_{j=1}^{\infty} \mathrm{x}_j\Big| \le \sum_{j=1}^{\infty} |\mathrm{x}_j|.$$

Proof Define

$$S_m^1 = \Big|\sum_{j=1}^{m} x_j\Big|, \quad S_m^2 = \sum_{j=1}^{m} |x_j|, \qquad m \in \mathbb{Z}_{>0}.$$

By Exercise 2.4.3 we have $S_m^1 \le S_m^2$ for each $m \in \mathbb{Z}_{>0}$. Moreover, by Proposition 2.4.3 the sequences $(S_m^1)_{m\in\mathbb{Z}_{>0}}$ and $(S_m^2)_{m\in\mathbb{Z}_{>0}}$ converge. It is then clear (why?) that

$$\lim_{m\to\infty} S_m^1 \le \lim_{m\to\infty} S_m^2,$$

which is the result. ■

It is not immediately clear on a first encounter why the notion of absolute convergence is useful. However, as we shall see in Chapter 8, it is the notion of absolute convergence that will be of most use to us in our characterisation of discrete signal spaces. The following result indicates why mere convergence of a series is perhaps not as nice a notion as one would like, and that absolute convergence is in some sense better behaved.***missing stuff***

2.4.5 Theorem (Convergence and rearrangement of series) *For a series* $\mathrm{S} = \sum_{j=1}^{\infty} \mathrm{x}_j$*, the following statements hold:*

(i) *if* S *is conditionally convergent then, for any* $\mathrm{s}_0 \in \mathbb{R}$*, there exists a bijection* $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ *such that the series* $\mathrm{S}_\phi = \sum_{j=1}^{\infty} \mathrm{x}_{\phi(j)}$ *converges to* s_0*;*

(ii) *if* S *is conditionally convergent then there exists a bijection* $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ *such that the series* $\mathrm{S}_\phi = \sum_{j=1}^{\infty} \mathrm{x}_{\phi(j)}$ *diverges to* ∞*;*

(iii) *if* S *is conditionally convergent then there exists a bijection* $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ *such that the series* $\mathrm{S}_\phi = \sum_{j=1}^{\infty} \mathrm{x}_{\phi(j)}$ *diverges to* $-\infty$*;*

(iv) *if* S *is conditionally convergent then there exists a bijection* $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ *such that the limit of the partial sums for the series* $\mathrm{S}_\phi = \sum_{j=1}^{\infty} \mathrm{x}_{\phi(j)}$ *is oscillating;*

(v) *if* S *is absolutely convergent then, for any bijection* $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$*, the series* $\mathrm{S}_\phi = \sum_{j=1}^{\infty} \mathrm{x}_{\phi(j)}$ *converges to the same limit as the series* S.

Proof We shall be fairly "descriptive" concerning the first four parts of the proof. More precise arguments can be tediously fabricated from the ideas given. We shall use the fact, given as Exercise 2.4.1, that if a series is conditionally convergent, then the two series formed by the positive terms and the negative terms diverge.

(i) First of all, rearrange the terms in the series so that the positive terms are arranged in decreasing order, and the negative terms are arranged in increasing order. We suppose that $s_0 \ge 0$, as a similar argument can be fabricated when $s_0 < 0$. Take as the first elements of the rearranged sequence the enough of the first few positive terms in the sequence so that their sum exceeds s_0. As the next terms, take enough of the first few negative terms in the series such that their sum, combined with the already chosen positive terms, is less than s_0. Now repeat this process. Because the series was initially rearranged so that the positive and negative terms are in descending and ascending order, respectively, one can show that the construction we have given yields a sequence of partial sums that starts greater than s_0, then monotonically decreases to a value less than s_0, then monotonically increases to a value greater than s_0, and so on. Moreover, at the end of each step, the values get closer to s_0 since the sequence of positive and negative terms both converge to zero. An argument like that used in the proof of Proposition 2.3.10 can then be used to show that the resulting sequence of partial sums converges to s_0.

(ii) To get the suitable rearrangement, proceed as follows. Partition the negative terms in the sequence into disjoint finite sets S^-_j, $j \in \mathbb{Z}_{>0}$. Now partition the positive terms in the sequence as follows. Define S^+_1 to be the first N_1 positive terms in the sequence, where N_1 is sufficiently large that the sum of the elements of S^+_1 exceeds by at least 1 in absolute value the sum of the elements from S^-_1. This is possible since the series of positive terms in the sequence diverges to ∞. Now define S^+_2 by taking the next N_2 positive terms in the sequence so that the sum of the elements of S^+_2 exceeds by at least 1 in absolute value the sum of the elements from S^-_2. Continue in this way, defining $S^+_3, S^+_4, \ldots$. The rearrangement of the terms in the series is then made by taking the first collection of terms to be the elements of S^+_1, the second collection to be the elements of S^-_1, the third collection to be the elements of S^+_2, and so on. One can verify that the resulting sequence of partial sums diverges to ∞.

(iii) The argument here is entirely similar to the previous case.

(iv) This result follows from part (i) in the following way. Choose an oscillating sequence $(y_j)_{j\in\mathbb{Z}_{>0}}$. For y_1, by part (i) one can find a finite number of terms from the original series whose sum is as close as desired to y_1. These will form the first terms in the rearranged series. Next, the same argument can be applied to the remaining elements of the series to yield a finite number of terms in the series that are as close as desired to y_2. One carries on in this way, noting that since the sequence $(y_j)_{j\in\mathbb{Z}_{>0}}$ is oscillating, so too will be the sequence of partial sums for the rearranged series.

(v) Let $y_j = x_{\phi(j)}$ for $j \in \mathbb{Z}_{>0}$. Then define sequences $(x^+_j)_{j\in\mathbb{Z}_{>0}}$, $(x^-_j)_{j\in\mathbb{Z}_{>0}}$, $(y^+_j)_{j\in\mathbb{Z}_{>0}}$, and $(y^-_j)_{j\in\mathbb{Z}_{>0}}$ by

$$x^+_j = \max\{x_j, 0\}, \quad x^-_j = \max\{-x_j, 0\},$$
$$y^+_j = \max\{y_j, 0\}, \quad y^-_j = \max\{-y_j, 0\}, \qquad j \in \mathbb{Z}_{>0},$$

noting that $|x_j| = \max\{x^-_j, x^+_j\}$ and $|y_j| = \max\{y^-_j, y^+_j\}$ for $j \in \mathbb{Z}_{>0}$. By Proposition 2.4.8 it follows that the series

$$S^+ = \sum_{j=1}^\infty x^+_j, \quad S^- = \sum_{j=1}^\infty x^-_j, \quad S^+_\phi = \sum_{j=1}^\infty y^+_j, \quad S^-_\phi = \sum_{j=1}^\infty y^-_j$$

converge. We claim that for each $k \in \mathbb{Z}_{>0}$ we have

$$\sum_{j=1}^k x^+_j \leq \sum_{j=1}^\infty y^+_j.$$

To see this, we need only note that there exists $N \in \mathbb{Z}_{>0}$ such that

$$\{x^+_1, \ldots, x^+_k\} \subseteq \{y^+_1, \ldots, y^+_N\}.$$

With N having this property,

$$\sum_{j=1}^{k} x_j^+ \leq \sum_{j=1}^{N} y_j^+ \leq \sum_{j=1}^{\infty} y_j^+,$$

as desired. Therefore,

$$\sum_{j=1}^{\infty} x_j^+ \leq \sum_{j=1}^{\infty} y_j^+.$$

Reversing the argument gives

$$\sum_{j=1}^{\infty} y_j^+ \leq \sum_{j=1}^{\infty} x_j^+ \quad \implies \quad \sum_{j=1}^{\infty} x_j^+ = \sum_{j=1}^{\infty} y_j^+.$$

A similar argument also gives

$$\sum_{j=1}^{\infty} x_j^- = \sum_{j=1}^{\infty} y_j^-.$$

This then gives

$$\sum_{j=1}^{\infty} y_j = \sum_{j=1}^{\infty} y_j^+ - \sum_{j=1}^{\infty} y_j^- = \sum_{j=1}^{\infty} x_j^+ - \sum_{j=1}^{\infty} x_j^- = \sum_{j=1}^{\infty} x_j,$$

as desired. ■

The theorem says, roughly, that absolute convergence is necessary and sufficient to ensure that the limit of a series be independent of rearrangement of the terms in the series. Note that the necessity portion of this statement, which is parts (i)–(iv) of the theorem, comes in a rather dramatic form which suggests that conditional convergence behaves maximally poorly with respect to rearrangement.

2.4.2 Tests for convergence of series

In this section we give some of the more popular tests for convergence of a series. It is infeasible to expect an easily checkable general condition for convergence. However, in some cases the tests we give here are sufficient.

First we make a simple general observation that is very often useful; it is merely a reflection that the convergence of a series depends only on the tail of the series. We shall often make use of this result without mention.

2.4.6 Proposition (Convergence is unaffected by changing a finite number of terms) *Let $\sum_{j=1} x_j$ and $\sum_{j=1}^{\infty} y_j$ be series in $\mathbb{R}$ and suppose that there exists $K \in \mathbb{Z}$ and $N \in \mathbb{Z}_{>0}$ such that $x_j = y_{j+K}$ for $j \geq N$. Then the following statements hold:*

(i) the series $\sum_{j=1}^{\infty} x_j$ converges if and only if the series $\sum_{j=1}^{\infty} y_j$ converges;

(ii) the series $\sum_{j=1}^{\infty} x_j$ diverges if and only if the series $\sum_{j=1}^{\infty} y_j$ diverges;

(iii) the series $\sum_{j=1}^{\infty} x_j$ diverges to ∞ if and only if the series $\sum_{j=1}^{\infty} y_j$ diverges to ∞;

(iv) the series $\sum_{j=1}^{\infty} x_j$ diverges to $-\infty$ if and only if the series $\sum_{j=1}^{\infty} y_j$ diverges to $-\infty$.

The next convergence result is also a more or less obvious one.

2.4.7 Proposition (Sufficient condition for a series to diverge) *If the sequence* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *does not converge to zero, then the series* $\sum_{j=1}^{\infty} x_j$ *diverges.*

Proof Suppose that the series $\sum_{j=1}^{\infty} x_j$ converges to s_0 and let $(S_k)_{k\in\mathbb{Z}_{>0}}$ be the sequence of partial sums. Then $x_k = S_k - S_{k-1}$. Then

$$\lim_{k\to\infty} x_k = \lim_{k\to\infty} S_k - \lim_{k\to\infty} S_{k-1} = s_0 - s_0 = 0_{\mathsf{V}},$$

as desired. ■

Note that Example 2.4.2–2 shows that the converse of this result is false. That is to say, for a series to converge, it is not sufficient that the terms in the series go to zero. For this reason, checking the convergence of a series numerically becomes something that must be done carefully, since the blind use of the computer with a prescribed numerical accuracy will suggest the false conclusion that a series converges if and only if the terms in the series go to zero as the index goes to infinity.

Another more or less obvious result is the following.

2.4.8 Proposition (Comparison Test) *Let* $(x_j)_{j\in\mathbb{Z}_{>0}}$ *and* $(y_j)_{j\in\mathbb{Z}_{>0}}$ *be sequences of nonnegative numbers for which there exists* $\alpha \in \mathbb{R}_{>0}$ *satisfying* $y_j \leq \alpha x_j$, $j \in \mathbb{Z}_{>0}$*. Then the following statements hold:*

(i) the series $\sum_{j=1}^{\infty} y_j$ *converges if the series* $\sum_{j=1}^{\infty} x_j$ *converges;*

(ii) the series $\sum_{j=1}^{\infty} x_j$ *diverges if the series* $\sum_{j=1}^{\infty} y_j$ *diverges.*

Proof We shall show that, if the series $\sum_{j=1}^{\infty} x_j$ converges, then the sequence $(T_k)_{k\in\mathbb{Z}_{>0}}$ of partial sums for the series $\sum_{j=1}^{\infty} y_j$ is a Cauchy sequence. Since the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ for $\sum_{j=1}^{\infty} x_j$ is convergent, it is Cauchy. Therefore, for $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that whenever $k, m \geq N$, with $k > m$ without loss of generality,

$$S_k - S_m = \sum_{j=m+1}^{k} x_j < \epsilon\alpha^{-1}.$$

Then, for $k, m \geq N$ with $k > m$ we have

$$T_k - T_m = \sum_{j=m+1}^{k} y_j \leq \alpha \sum_{j=m+1}^{k} x_j < \epsilon,$$

showing that $(T_k)_{k\in\mathbb{Z}_{>0}}$ is a Cauchy sequence, as desired.

The second statement is the contrapositive of the first. ■

Now we can get to some less obvious results for convergence of series. The first result concerns series where the terms alternate sign.

2.4.9 Proposition (Alternating Test) *Let* $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathbb{R}$ *satisfying*

(i) $\mathrm{x}_j > 0$ *for* $j \in \mathbb{Z}_{>0}$*,*

(ii) $\mathrm{x}_{j+1} \leq \mathrm{x}_j$ *for* $j \in \mathbb{Z}_{>0}$*, and*

(iii) $\lim_{j\to\infty} \mathrm{x}_j = 0$.

Then the series $\sum_{j=1}^{\infty}(-1)^{j+1}\mathrm{x}_j$ *converges.*

Proof The proof is a straightforward generalisation of that given for Example 2.4.2–3, and we leave for the reader the simple exercise of verifying that this is so. ■

Our next result is one that is often useful.

2.4.10 Proposition (Ratio Test for series) *Let* $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ *be a nonzero sequence in* $\mathbb{R}$ *with* $\sum_{j=1}^{\infty} \mathrm{x}_j$ *the corresponding series. Then the following statements hold:*

(i) if $\limsup_{j\to\infty}\left|\frac{\mathrm{x}_{j+1}}{\mathrm{x}_j}\right| < 1$*, then the series converges absolutely;*

(ii) if there exists $N \in \mathbb{Z}_{>0}$ *such that* $\left|\frac{\mathrm{x}_{j+1}}{\mathrm{x}_j}\right| > 1$ *for all* $j \geq N$*, then the series diverges.*

Proof (i) By Proposition 2.3.15 there exists $\beta \in (0,1)$ and $N \in \mathbb{Z}_{>0}$ such that $\left|\frac{x_{j+1}}{x_j}\right| < \beta$ for $j \geq N$. Then

$$\left|\frac{x_j}{x_N}\right| = \left|\frac{x_{N+1}}{x_N}\right|\left|\frac{x_{N+2}}{x_{N+1}}\right|\cdots\left|\frac{x_j}{x_{j-1}}\right| < \beta^{j-N}, \qquad j > N,$$

implying that

$$|x_j| < \frac{|x_N|}{\beta^N}\beta^j.$$

Since $\beta < 1$, the geometric series $\sum_{j=1}^{\infty}\beta^j$ converges. The result for $\alpha < 1$ now follows by the Comparison Test.

(ii) The sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ cannot converge to 0 in this case, and so this part of the result follows from Proposition 2.4.7. ■

The following simpler test is often stated as the Ratio Test.

2.4.11 Corollary (Weaker version of the Ratio Test) *If* $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ *is a nonzero sequence in* $\mathbb{R}$ *for which* $\lim_{j\to\infty}\left|\frac{\mathrm{x}_{j+1}}{\mathrm{x}_j}\right| = \alpha$*, then the series* $\sum_{j=1}^{\infty}\mathrm{x}_j$ *converges absolutely if* $\alpha < 1$ *and diverges if* $\alpha > 1$.

2.4.12 Remark (Nonzero assumption in Ratio Test) In the preceding two results we asked that the terms in the series be nonzero. This is not a significant limitation. Indeed, one can enumerate the nonzero terms in the series, and then apply the ratio test to this. •

Our next result has a similar character to the previous one.

2.4.13 Proposition (Root Test) *Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence for which $\limsup_{j\to\infty}|x_j|^{1/j} = \alpha$. Then the series $\sum_{j=1}^{\infty} x_j$ converges absolutely if $\alpha < 1$ and diverges if $\alpha > 1$.*

Proof First take $\alpha < 1$ and define $\beta = \frac{1}{2}(\alpha + 1)$. Then, just as in the proof of Proposition 2.4.10, $\alpha < \beta < 1$. By Proposition 2.3.15 there exists $N \in \mathbb{Z}_{>0}$ such that $|x_j|^{1/j} < \beta$ for $j \geq N$. Thus $|x_j| < \beta^j$ for $j \geq N$. Note that $\sum_{j=N+1}^{\infty} \beta^j$ converges by Example 2.4.2–1. Now $\sum_{j=0}^{\infty}|x_j|$ converges by the Comparison Test.

Next take $\alpha > 1$. In this case we have $\lim_{j\to\infty}|x_j| \neq 0$, and so we conclude divergence from Proposition 2.4.7. ■

The following obvious corollary is often stated as the Root Test.

2.4.14 Corollary (Weaker version of Root Test) *Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence for which $\lim_{j\to\infty}|x_j|^{1/j} = \alpha$. Then the series $\sum_{j=1}^{\infty} x_j$ converges absolutely if $\alpha < 1$ and diverges if $\alpha > 1$.*

The Ratio Test and the Root Test are related, as the following result indicates.

2.4.15 Proposition (Root Test implies Ratio Test) *If $(p_j)_{j\in\mathbb{Z}_{\geq 0}}$ is a sequence in $\mathbb{R}_{>0}$ then*

$$\liminf_{j\to\infty} \frac{p_{j+1}}{p_j} \leq \liminf_{j\to\infty} p_j^{1/j}$$
$$\limsup_{j\to\infty} p_j^{1/j} \leq \limsup_{j\to\infty} \frac{p_{j+1}}{p_j}.$$

In particular, for a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$, if $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right|$ exists, then $\lim_{j\to\infty}|x_j|^{1/j} = \lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right|$.

Proof For the first inequality, let $\alpha = \liminf_{j\to\infty} \frac{p_{j+1}}{p_j}$. First consider the case where $\alpha = \infty$. Then, given $M \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $\frac{p_{j+1}}{p_j} > M$ for $j \geq N$. Then we have

$$\left|\frac{p_j}{p_N}\right| = \left|\frac{p_{N+1}}{p_N}\right|\left|\frac{p_{N+1}}{p_{N+1}}\right| \cdots \left|\frac{p_j}{p_{j-1}}\right| > M^{j-N}, \qquad j > N.$$

This gives

$$p_j > \frac{p_N}{M^N} M^j, \qquad j > N.$$

Thus $p_j^{1/j} > (\frac{p_N}{M^N})^{1/j} M$. Since $\lim_{j\to\infty}(p_N\beta^{-N})^{1/j} = 1$ (cf. the definition of P_a in Section 3.6.3), we have $\liminf_{j\to\infty} p_j^{1/j} > M$, giving the desired conclusion in this case, since M is arbitrary. Next consider the case when $\alpha \in \mathbb{R}_{>0}$ and let $\beta < \alpha$. By Proposition 2.3.16 there exists $N \in \mathbb{Z}_{>0}$ such that $\frac{p_{j+1}}{p_j} \geq \beta$ for $j \geq N$. Performing just the same computation as above gives $p_j \geq \beta^{j-N} p_N$ for $j \geq N$. Therefore, $p_j^{1/j} \geq (p_N\beta^{-N})^{1/j}\beta$. Since $\lim_{j\to\infty}(p_N\beta^{-N})^{1/j} = 1$ we have $\liminf_{j\to\infty} p_j^{1/j} \geq \beta$. The first inequality follows since $\beta < \alpha$ is arbitrary.

Now we prove the second inequality. Let $\alpha = \limsup_{j\to\infty} \frac{p_{j+1}}{p_j}$. If $\alpha = \infty$ then the second inequality in the statement of the result is trivial. If $\alpha \in \mathbb{R}_{>0}$ then let $\beta > \alpha$ and note that there exists $N \in \mathbb{Z}_{>0}$ such that $\frac{p_{j+1}}{p_j} \leq \beta$ for $j \geq N$ by Proposition 2.3.15. In particular, just as in the proof of Proposition 2.4.10, $p_j \leq \beta^{j-N} p_N$ for $j \geq N$. Therefore,

$p_j^{1/j} \leq (p_N\beta^{-N})^{1/j}\beta$. Since $\lim_{j\to\infty}(p_N\beta^{-N})^{1/j} = 1$ we then have $\liminf_{j\to\infty} p_j^{1/j} \leq \beta$. the second inequality follows since $\beta > \alpha$ is arbitrary.

The final assertion follows immediately from the two inequalities using Proposition 2.3.17. ■

In Exercises 2.4.6 and 2.4.7 the reader can explore the various possibilities for the ratio test and root test when $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and $\lim_{j\to\infty}|x_j|^{1/j} = 1$, respectively.

The final result we state in this section can be thought of as the summation version of integration by parts.

2.4.16 Proposition (Abel's[6] partial summation formula) *For sequences* $(\mathrm{x}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(\mathrm{y}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *of real numbers, denote* $\mathrm{S}_\mathrm{k} = \sum_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{x}_\mathrm{j}$. *Then*

$$\sum_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{x}_\mathrm{j}\mathrm{y}_\mathrm{j} = \mathrm{S}_\mathrm{k}\mathrm{y}_{\mathrm{k}+1} - \sum_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{S}_\mathrm{j}(\mathrm{y}_{\mathrm{j}+1} - \mathrm{y}_\mathrm{j}) = \mathrm{S}_\mathrm{k}\mathrm{y}_1 + \sum_{\mathrm{j}=1}^{\mathrm{k}}(\mathrm{S}_\mathrm{k} - \mathrm{S}_\mathrm{j})(\mathrm{y}_{\mathrm{j}+1} - \mathrm{y}_\mathrm{j}).$$

Proof Let $S_0 = 0$ by convention. Since $x_j = S_j - S_{j-1}$ we have

$$\sum_{j=1}^{n} x_j y_j = \sum_{j=1}^{n}(S_j - S_{j-1})y_j = \sum_{j=1}^{n} S_j y_j - \sum_{j=1}^{n} S_{j-1}y_j.$$

Trivially,

$$\sum_{j=1}^{n} S_{j-1}y_j = \sum_{j=1}^{n} S_j y_{j+1} - S_n y_{n+1}.$$

This gives the first equality of the lemma. The second follows from a substitution of

$$y_{n+1} = \sum_{j=1}^{n}(y_{j+1} - y_j) + y_1$$

into the first equality. ■

2.4.3 e and π

In this section we consider two particular convergent series whose limits are among the most important of "physical constants."

2.4.17 Definition (e) $\mathrm{e} = \displaystyle\sum_{j=0}^{\infty}\frac{1}{j!}$. •

Note that the series defining e indeed converges, for example, by the Ratio Test. Another common representation of e as a limit is the following.

[6] Niels Henrik Abel (1802–1829) was a Norwegian mathematician who worked in the area of analysis. An important theorem of Abel, one that is worth knowing for people working in application areas, is a theorem stating that there is no expression for the roots of a quintic polynomial in terms of the coefficients that involves only the operations of addition, subtraction, multiplication, division and taking roots.

2.4.18 Proposition (Alternative representations of e) *We have*

$$\mathrm{e} = \lim_{\mathrm{j}\to\infty}\left(1+\tfrac{1}{\mathrm{j}}\right)^{\mathrm{j}} = \lim_{\mathrm{j}\to\infty}\left(1+\tfrac{1}{\mathrm{j}}\right)^{\mathrm{j}+1}.$$

Proof First note that if the limit $\lim_{j\to\infty}\left(1+\frac{1}{j}\right)^{j}$ exists, then, by Proposition 2.3.23,

$$\lim_{j\to\infty}\left(1+\tfrac{1}{j}\right)^{j+1} = \lim_{j\to\infty}\left(1+\tfrac{1}{j}\right)\left(1+\tfrac{1}{j}\right)^{j} = \lim_{j\to\infty}\left(1+\tfrac{1}{j}\right)^{j}.$$

Thus we will only prove that $\mathrm{e} = \lim_{j\to\infty}\left(1+\frac{1}{j}\right)^{j}$.

Let

$$S_k = \sum_{j=0}^{k}\frac{1}{k!},\quad A_k = \left(1+\tfrac{1}{k}\right)^{k},\quad B_k = \left(1+\tfrac{1}{k}\right)^{k+1},$$

be the kth partial sum of the series for e and the kth term in the proposed sequence for e. By the Binomial Theorem (Exercise 2.2.1) we have

$$A_k = \left(1+\tfrac{1}{k}\right)^{k} = \sum_{j=0}^{k}\binom{k}{j}\frac{1}{k^j}.$$

Moreover, the exact form for the binomial coefficients can directly be seen to give

$$A_k = \sum_{j=0}^{k}\frac{1}{j!}\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\cdots\left(1-\frac{j-1}{k}\right).$$

Each coefficient of $\frac{1}{j!}$, $j\in\{0,1,\dots,k\}$ is then less than 1. Thus $A_k \le S_k$ for each $k\in\mathbb{Z}_{\ge 0}$. Therefore, $\limsup_{k\to\infty} A_k \le \limsup_{k\to\infty} S_k$. For $m\le k$ the same computation gives

$$A_k \ge \sum_{j=0}^{m}\frac{1}{j!}\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\cdots\left(1-\frac{j-1}{k}\right).$$

Fixing m and letting $k\to\infty$ gives

$$\liminf_{k\to\infty} A_k \ge \sum_{j=0}^{m}\frac{1}{j!} = S_m.$$

Thus $\liminf_{k\to\infty} A_k \ge \liminf_{m\to\infty} S_m$, which gives the result when combined with our previous estimate $\limsup_{k\to\infty} A_k \le \limsup_{k\to\infty} S_k$. ■

It is interesting to note that the series representation of e allows us to conclude that e is irrational.

2.4.19 Proposition (Irrationality of e) $\mathrm{e} \in \mathbb{R} \setminus \mathbb{Q}$.

Proof Suppose that $\mathrm{e} = \frac{l}{m}$ for $l, m \in \mathbb{Z}_{>0}$. We compute

$$(m-1)!l = m!\mathrm{e} = m! \sum_{j=0}^{\infty} \frac{1}{j!} = \sum_{j=0}^{m} \frac{m!}{j!} + \sum_{j=m+1}^{\infty} \frac{m!}{j!},$$

which then gives

$$\sum_{j=m+1}^{\infty} \frac{m!}{j!} = (m-1)!l - \sum_{j=0}^{m} \frac{m!}{j!},$$

which implies that $\sum_{j=m+1}^{\infty} \frac{m!}{j!} \in \mathbb{Z}_{>0}$. We then compute, using Example 2.4.2–1,

$$0 < \sum_{j=m+1}^{\infty} \frac{m!}{j!} < \sum_{j=m+1}^{\infty} \frac{1}{(m+1)^{j-m}} = \sum_{j=1}^{\infty} \frac{1}{(m+1)^{j}} = \frac{\frac{1}{m+1}}{1 - \frac{1}{m+1}} = \frac{1}{m} \leq 1.$$

Thus $\sum_{j=m+1}^{\infty} \frac{m!}{j!} \in \mathbb{Z}_{>0}$, being an integer, must equal 1, and, moreover, $m = 1$. Thus we have

$$\sum_{j=2}^{\infty} \frac{1}{j!} = \mathrm{e} - 2 = 1 \quad \implies \quad \mathrm{e} = 3.$$

Next let

$$\alpha = \sum_{j=1}^{\infty} \Big(\frac{1}{2^{j-1}} - \frac{1}{j!} \Big),$$

noting that this series for α converges, and converges to a positive number since each term in the series is positive. Then, using Example 2.4.2–1,

$$\alpha = (2 - (\mathrm{e} - 1)) \quad \implies \quad \mathrm{e} = 3 - \alpha.$$

Thus $\mathrm{e} < 3$, and we have arrived at a contradiction. ■

Next we turn to the number π. Perhaps the best description of π is that it is the ratio of the circumference of a circle with the diameter of the circle. Indeed, the use of the Greek letter "p" (i.e., π) has its origins in the word "perimeter." However, to make sense of this definition, one must be able to talk effectively about circles, what the circumference means, etc. This is more trouble than it is worth for us at this point. Therefore, we give a more analytic description of π, albeit one that, at this point, is not very revealing of what the reader probably already knows about it.

2.4.20 Definition (π) $\pi = 4 \sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1}$. •

By the Alternating Test, this series representation for π converges.

We can also fairly easily show that π is irrational, although our proof uses some facts about functions on $\mathbb{R}$ that we will not discuss until Chapter 3.

2.4.21 Proposition (Irrationality of π) $\pi \in \mathbb{R} \setminus \mathbb{Q}$.

Proof In Section 3.6.4 we will give a definition of the trigonometric functions, sin and cos, and prove that, on $(0, \pi)$, sin is positive, and that $\sin 0 = \sin \pi = 0$. We will also prove the rules of differentiation for trigonometric functions necessary for the proof we now present.

Note that if π is rational, then π^2 is also rational. Therefore, it suffices to show that π^2 is irrational.

Let us suppose that $\pi^2 = \frac{l}{m}$ for $l, m \in \mathbb{Z}_{>0}$. For $k \in \mathbb{Z}_{>0}$ define $f_k\colon [0,1] \to \mathbb{R}$ by

$$f_k(x) = \frac{x^k(1-x)^k}{k!},$$

noting that $\operatorname{image}(f) \subseteq [0, \frac{1}{k!}]$. It is also useful to write

$$f_k(x) = \frac{1}{k!}\sum_{j=k}^{2k} c_j x^j,$$

where we observe that c_j, $j \in \{k, k+1, \dots, 2k\}$ are integers. Define $g_j\colon [0,1] \to \mathbb{R}$ by

$$g_k(x) = k^j \sum_{j=0}^{k} (-1)^j \pi^{2(k-j)} f^{(2j)}(x).$$

A direct computation shows that

$$f_k^{(j)}(0) = 0, \qquad j < k,\ j > 2k,$$

and that

$$f_k^{(j)}(0) = \frac{j!}{k!}c_j, \qquad j \in \{k, k+1, \dots, 2k\},$$

is an integer. Thus f and all of its derivatives take integer values at $x = 0$, and therefore also at $x = 1$ since $f_k(x) = f_k(1-x)$. One also verifies directly that $g_k(0)$ and $g_k(1)$ are integers.

Now we compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}(g_k'(x)\sin \pi x - \pi g_k(x)\cos \pi x) &= (g_k''(x) + \pi^2 g_k(x))\sin \pi x \\ &= m^k \pi^{2k+2} f(x) \sin \pi x = \pi^2 l^k f(x) \sin \pi x,\end{aligned}$$

using the definition of g_k and the fact that $\pi^2 = \frac{l}{m}$. By the Fundamental Theorem of Calculus we then have, after a calculation,

$$\pi l^k \int_0^1 f(x) \sin \pi x \,\mathrm{d}x = g_k(0) + g_k(1) \in \mathbb{Z}_{>0}.$$

But we then have, since the integrand in the above integral is nonnegative,

$$0 < \pi l^k \int_0^1 f(x) \sin \pi x \,\mathrm{d}x < \frac{\pi l^k}{k!}$$

given the bounds on f_k. Note that $\lim_{k\to\infty} \frac{l^k}{k!} = 0$. Since the above computations hold for any k, if we take k sufficiently large that $\frac{\pi l^k}{k!} < 1$, we arrive at a contradiction. ■

2.4.4 Doubly infinite series

We shall frequently encounter series whose summation index runs not from 1 to ∞, but from $-\infty$ to ∞. Thus we call a family $(x_j)_{j\in\mathbb{Z}}$ of elements of $\mathbb{R}$ a ***doubly infinite sequence*** in $\mathbb{R}$, and a sum of the form $\sum_{j=-\infty}^{\infty} x_j$ a ***doubly infinite series***. A little care need to be shown when defining convergence for such series, and here we give the appropriate definitions.

2.4.22 Definition (Convergence and absolute convergence of doubly infinite series) Let $(x_j)_{j\in\mathbb{Z}}$ be a doubly infinite sequence and let $S = \sum_{j=-\infty}^{\infty} x_j$ be the corresponding doubly infinite series. The sequence of ***single partial sums*** is the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ where

$$S_k = \sum_{j=-k}^{k} x_j,$$

and the sequence of ***double partial sums*** is the double sequence $(S_{k,l})_{k,l\in\mathbb{Z}_{>0}}$ defined by

$$S_{k,l} = \sum_{j=-k}^{l} x_j.$$

Let $s_0 \in \mathbb{R}$. The doubly infinite series:

(i) ***converges to*** $\mathbf{s_0}$ if the double sequence of partial sums converges to s_0;

(ii) has s_0 as a ***limit*** if it converges to s_0;

(iii) is ***convergent*** if it converges to some element of $\mathbb{R}$;

(iv) ***converges absolutely***, or is ***absolutely convergent***, if the doubly infinite series

$$\sum_{j=-\infty}^{\infty} |x_j|$$

converges;

(v) ***converges conditionally***, or is ***conditionally convergent***, if it is convergent, but not absolutely convergent;

(vi) ***diverges*** if it does not converge;

(vii) ***diverges to*** $\boldsymbol{\infty}$ (resp. ***diverges to*** $\boldsymbol{-\infty}$), and we write $\sum_{j=-\infty}^{\infty} x_j = \infty$ (resp. $\sum_{j=-\infty}^{\infty} x_j = -\infty$), if the sequence of double partial sums diverges to ∞ (resp. diverges to $-\infty$);

(viii) has a limit that ***exists*** if $\sum_{j=-\infty}^{\infty} x_j \in \mathbb{R}$;

(ix) is ***oscillatory*** if the limit of the double sequence of partial sums is oscillatory. •

2.4.23 Remark (Partial sums versus double partial sums) Note that the convergence of the sequence of partial sums is not a very helpful notion, in general. For example, the series $\sum_{j=-\infty}^{\infty} j$ possesses a sequence of partial sums that is identically zero, and so the sequence of partial sums obviously converges to zero. However, it is not likely that one would wish this doubly infinite series to qualify as convergent. Thus partial sums are not a particularly good measure of convergence. However, there are situations—for example, the convergence of Fourier series (see Chapter 12)—where the standard notion of convergence of a doubly infinite series is made using the partial sums. However, in these cases, there is additional structure on the setup that makes this a reasonable thing to do. •

The convergence of a doubly infinite series has the following useful, intuitive characterisation.

2.4.24 Proposition (Characterisation of convergence of doubly infinite series) *For a doubly infinite series* $\mathrm{S} = \sum_{\mathrm{j}=-\infty}^{\infty} \mathrm{x}_{\mathrm{j}}$*, the following statements are equivalent:*

(i) S *converges;*

(ii) the two series $\sum_{\mathrm{j}=0}^{\infty} \mathrm{x}_{\mathrm{j}}$ *and* $\sum_{\mathrm{j}=1}^{\infty} \mathrm{x}_{-\mathrm{j}}$ *converge.*

Proof For $k, l \in \mathbb{Z}_{>0}$, denote

$$S_{k,l} = \sum_{-k}^{l} x_j, \quad S_k^+ = \sum_{j=0}^{k} x_j, \quad S_k^- = \sum_{-k}^{-1} x_j,$$

so that $S_{k,l} = S_k^- + S_l^+$.

(i) $\Longrightarrow$ (ii) Let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $|S_{j,k} - s_0| < \frac{\epsilon}{2}$ for $j, k \geq N$. Now let $j, k \geq N$, choose some $l \geq N$, and compute

$$|S_j^+ - S_k^+| \leq |S_j^+ + S_l^- - s_0| + |S_k^+ + S_l^- - s_0| < \epsilon.$$

Thus $(S_j^+)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence, and so is convergent. A similar argument shows that $(S_j^-)_{j\in\mathbb{Z}_{>0}}$ is also a Cauchy sequence.

(ii) $\Longrightarrow$ (i) Let s^+ be the limit of $\sum_{j=0}^{\infty} x_j$ and let s^- be the limit of $\sum_{j=1}^{\infty} x_{-j}$. For $\epsilon \in \mathbb{R}_{>0}$ define $N^+, N^- \in \mathbb{Z}_{>0}$ such that $|S_j^+ - s^+| < \frac{\epsilon}{2}$, $j \geq N^+$, and $|S_j^- - s^-| < \frac{\epsilon}{2}$, $j \leq -N^-$. Then, for $j, k \geq \max\{N^-, N^+\}$,

$$|S_{j,k} - (s^+ + s^-)| = |S_k^+ - s_+ + S_j^- - s_-| \leq |S_k^+ - s_+| + |S_j^- - s_-| < \epsilon,$$

thus showing that S converges. ■

Thus convergent doubly infinite series are really just combinations of convergent series in the sense that we have studied in the preceding sections. Thus, for example, one can use the tests of Section 2.4.2 to check for convergence of a doubly infinite series by applying them to both "halves" of the series. Also, the relationships between convergence and absolute convergence for series also hold for doubly infinite series. And a suitable version of Theorem 2.4.5 also holds for doubly infinite series. These facts are so straightforward that we will assume them in the sequel without explicit mention; they all follow directly from Proposition 2.4.24.

2.4.5 Multiple series

Just as we considered multiple sequences in Section 2.3.5, we can consider multiple series. As we did with sequences, we content ourselves with double series.

2.4.25 Definition (Double series) A ***double series*** in $\mathbb{R}$ is a sum of the form $\sum_{j,k=1}^{\infty} x_{jk}$ where $(x_{jk})_{j,k\in\mathbb{Z}_{>0}}$ is a double sequence in $\mathbb{R}$. •

While our definition of a series was not entirely sensible since it was not really identifiable as anything unless it had certain convergence properties, for double series, things are even worse. In particular, it is not clear what $\sum_{j,k=1}^{\infty} x_{jk}$ means. Does it mean $\sum_{j=1}^{\infty}\big(\sum_{k=1}^{\infty} x_{jk}\big)$? Does it mean $\sum_{k=1}^{\infty}\big(\sum_{j=1}^{\infty} x_{jk}\big)$? Or does it mean something different from both of these? The only way to rectify our poor mathematical manners is to define convergence for double series as quickly as possible.

2.4.26 Definition (Convergence and absolute convergence of double series) Let $(x_{jk})_{j,k\in\mathbb{Z}_{>0}}$ be a double sequence in $\mathbb{R}$ and consider the double series

$$S = \sum_{j,k=1}^{\infty} x_{jk}.$$

The corresponding sequence of ***partial sums*** is the double sequence $(S_{jk})_{j,k\in\mathbb{Z}_{>0}}$ defined by

$$S_{jk} = \sum_{l=1}^{j}\sum_{m=1}^{k} x_{lm}.$$

Let $s_0 \in \mathbb{R}$. The double series:

(i) ***converges to*** $\mathbf{s_0}$, and we write $\sum_{j,k=1}^{\infty} x_{jk} = s_0$, if the double sequence of partial sums converges to s_0;

(ii) has s_0 as a ***limit*** if it converges to s_0;

(iii) is ***convergent*** if it converges to some member of $\mathbb{R}$;

(iv) ***converges absolutely***, or is ***absolutely convergent***, if the series

$$\sum_{j,k=1}^{\infty} |x_{jk}|$$

converges;

(v) ***converges conditionally***, or is ***conditionally convergent***, if it is convergent, but not absolutely convergent;

(vi) ***diverges*** if it does not converge;

(vii) ***diverges to*** $\boldsymbol{\infty}$ (resp. ***diverges to*** $\boldsymbol{-\infty}$), and we write $\sum_{j,k=1}^{\infty} x_{jk} = \infty$ (resp. $\sum_{j,k=1}^{\infty} x_{jk} = -\infty$), if the double sequence of partial sums diverges to ∞ (resp. diverges to $-\infty$);

(viii) has a limit that ***exists*** if $\sum_{j,k=1}^{\infty} x_{jk} \in \mathbb{R}$;

(ix) is ***oscillatory*** if the sequence of partial sums is oscillatory. •

Note that the definition of the partial sums, S_{jk}, $j,k \in \mathbb{Z}_{>0}$, for a double series is unambiguous since

$$\sum_{l=1}^{j}\sum_{m=1}^{k} x_{lm} = \sum_{m=1}^{k}\sum_{l=1}^{j} x_{lm},$$

this being valid for finite sums. The idea behind convergence of double series, then, has an interpretation that can be gleaned from that in Figure 2.2 for double sequences.

Let us state a result, derived from similar results for double sequences, that allows the computation of limits of double series by computing one limit at a time.

2.4.27 Proposition (Computation of limits of double series I) *Suppose that for the double series $\sum_{\mathrm{j,k}=1}^{\infty} \mathrm{x_{jk}}$ it holds that*

(i) the double series is convergent and

(ii) for each $\mathrm{j} \in \mathbb{Z}_{>0}$, the series $\sum_{\mathrm{k}=1}^{\infty} \mathrm{x_{jk}}$ converges.

Then the series $\sum_{\mathrm{j}=1}^{\infty}(\sum_{\mathrm{k}=1}^{\infty} \mathrm{x_{jk}})$ converges and its limit is equal to $\sum_{\mathrm{j,k}=1}^{\infty} \mathrm{x_{jk}}$.

Proof This follows directly from Proposition 2.3.20. ■

2.4.28 Proposition (Computation of limits of double series II) *Suppose that for the double series $\sum_{\mathrm{j,k}=1}^{\infty} \mathrm{x_{jk}}$ it holds that*

(i) the double series is convergent,

(ii) for each $\mathrm{j} \in \mathbb{Z}_{>0}$, the series $\sum_{\mathrm{k}=1}^{\infty} \mathrm{x_{jk}}$ converges, and

(iii) for each $\mathrm{k} \in \mathbb{Z}_{>0}$, the limit $\sum_{\mathrm{j}=1}^{\infty} \mathrm{x_{jk}}$ converges.

Then the series $\sum_{\mathrm{j}=1}^{\infty}(\sum_{\mathrm{k}=1}^{\infty} \mathrm{x_{jk}})$ and $\sum_{\mathrm{k}=1}^{\infty}(\sum_{\mathrm{j}=1}^{\infty} \mathrm{x_{jk}})$ converge and their limits are both equal to $\sum_{\mathrm{j,k}=1}^{\infty} \mathrm{x_{jk}}$.

Proof This follows directly from Proposition 2.3.21. ■

missing stuff

2.4.6 Algebraic operations on series

In this section we consider the manner in which series interact with algebraic operations. The results here mirror, to some extent, the results for sequences in Section 2.3.6. However, the series structure allows for different ways of thinking about the product of sequences. Let us first give these definitions. For notational convenience, we use sums that begin at 0 rather than 1. This clearly has no affect on the definition of a series, or on any of its properties.

2.4.29 Definition (Products of series) Let $S = \sum_{j=0}^{\infty} x_j$ and $T = \sum_{j=0}^{\infty} y_j$ be series in $\mathbb{R}$.

(i) The ***product*** of S and T is the double series $\sum_{j,k=0}^{\infty} x_j y_k$.

(ii) The ***Cauchy product*** of S and T is the series $\sum_{k=0}^{\infty}\bigl(\sum_{j=0}^{k} x_j y_{k-j}\bigr)$. •

Now we can state the basic results on algebraic manipulation of series.

2.4.30 Proposition (Algebraic operations on series) *Let* $\mathrm{S} = \sum_{\mathrm{j}=0}^{\infty} \mathrm{x_j}$ *and* $\mathrm{T} = \sum_{\mathrm{j}=0}^{\infty} \mathrm{y_j}$ *be series in* $\mathbb{R}$ *that converges to* $\mathrm{s_0}$ *and* $\mathrm{t_0}$*, respectively, and let* $\alpha \in \mathbb{R}$*. Then the following statements hold:*

(i) the series $\sum_{\mathrm{j}=0}^{\infty} \alpha\mathrm{x_j}$ *converges to* $\alpha\mathrm{s_0}$*;*

(ii) the series $\sum_{\mathrm{j}=0}^{\infty}(\mathrm{x_j} + \mathrm{y_j})$ *converges to* $\mathrm{s_0} + \mathrm{t_0}$*;*

(iii) if S *and* T *are absolutely convergent, then the product of* S *and* T *is absolutely convergent and converges to* $\mathrm{s_0 t_0}$*;*

(iv) if S *and* T *are absolutely convergent, then the Cauchy product of* S *and* T *is absolutely convergent and converges to* $\mathrm{s_0 t_0}$*;*

(v) if S *or* T *are absolutely convergent, then the Cauchy product of* S *and* T *is convergent and converges to* $\mathrm{s_0 t_0}$*;*

(vi) if S *and* T *are convergent, and if the Cauchy product of* S *and* T *is convergent, then the Cauchy product of* S *and* T *converges to* $\mathrm{s_0 t_0}$*.*

Proof (i) Since $\sum_{j=0}^{k} \alpha x_j = \alpha \sum_{j=0}^{k} x_j$, this follows from part (i) of Proposition 2.3.23.

(ii) Since $\sum_{j=0}^{\infty}(x_j + y_j) = \sum_{j=0}^{k} x_j + \sum_{j=0}^{k} y_j$, this follows from part (ii) of Proposition 2.3.23.

(iii) and (iv) To prove these parts of the result, we first make a general argument. We note that $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ is a countable set (e.g., by Proposition **??**), and so there exists a bijection, in fact many bijections, $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$. For such a bijection ϕ, suppose that we are given a double sequence $(x_{jk})_{j,k\in\mathbb{Z}_{\geq 0}}$ and define a sequence $(x_j^{\phi})_{j\in\mathbb{Z}_{>0}}$ by $x_j^{\phi} = x_{kl}$ where $(k, l) = \phi(j)$. We then claim that, for any bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0}$, the double series $A = \sum_{k,l=1}^{\infty} x_{kl}$ converges absolutely if and only if the series $A^{\phi} = \sum_{j=1}^{\infty} x_j^{\phi}$ converges absolutely.

Indeed, suppose that the double series $|A| = \sum_{k,l=1}^{\infty}|x_{kl}|$ converges to $\beta \in \mathbb{R}$. For $\epsilon \in \mathbb{R}_{>0}$ the set

$$\{(k, l) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \mid \ ||A|_{kl} - \beta| \geq \epsilon\}$$

is then finite. Therefore, there exists $N \in \mathbb{Z}_{>0}$ such that, if $(k, l) = \phi(j)$ for $j \geq N$, then $||A|_{kl} - \beta| < \epsilon$. It therefore follows that $||A^{\phi}|_j - \beta| < \epsilon$ for $j \geq N$, where $|A^{\phi}|$ denotes the series $\sum_{j=1}^{\infty}|x_j^{\phi}|$. This shows that the series $|A^{\phi}|$ converges to β.

For the converse, suppose that the series $|A^{\phi}|$ converges to β. Then, for $\epsilon \in \mathbb{R}_{>0}$ the set

$$\{j \in \mathbb{Z}_{>0} \mid \ ||A^{\phi}|_j - \beta| \geq \epsilon\}$$

is finite. Therefore, there exists $N \in \mathbb{Z}_{>0}$ such that

$$\{(k, l) \in \mathbb{Z}_{\geq 0} \mid \ k, l \geq N\} \cap \{(k, l) \in \mathbb{Z}_{\geq 0} \mid \ ||A^{\phi}|_{\phi^{-1}(k,l)} - \beta| \geq \epsilon\} = \emptyset.$$

It then follows that for $k, l \geq N$ we have $||A|_{kl} - \beta| < \epsilon$, showing that $|A|$ converges to β.

Thus we have shown that A is absolutely convergent if and only if A^ϕ is absolutely convergent for any bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$. From part (v) of Theorem 2.4.5, and its generalisation to double series, we know that the limit of an absolutely convergent series or double series is independent of the manner in which the terms in the series are arranged.

Consider now a term in the product of S and T. It is easy to see that this term appears exactly once in the Cauchy product of S and T. Conversely, each term in the Cauchy product appears exactly one in the product. Thus the product and Cauchy product are simply rearrangements of one another. Moreover, each term in the product and the Cauchy product appears exactly once in the expression

$$\Big(\sum_{j=0}^{N} x_j\Big)\Big(\sum_{k=0}^{N} y_k\Big)$$

as we allow N to go to ∞. That is to say,

$$\sum_{j,k=0}^{\infty} x_j y_k = \sum_{k=0}^{\infty}\Big(\sum_{j=k}^{k} x_j y_{k-j}\Big) = \lim_{N\to\infty}\Big(\sum_{j=0}^{N} x_j\Big)\Big(\sum_{k=0}^{N} y_k\Big).$$

However, this last limit is exactly $s_0 t_0$, using part (iii) of Proposition 2.3.23.

(v) Without loss of generality, suppose that S converges absolutely. Let $(S_k)_{k\in\mathbb{Z}_{>0}}$, $(T_k)_{k\in\mathbb{Z}_{>0}}$, and $((ST)_k)_{k\in\mathbb{Z}_{>0}}$ be the sequences of partial sums for S, T, and the Cauchy product, respectively. Also define $\tau_k = T_k - t_0$, $k \in \mathbb{Z}_{\geq 0}$. Then

$$\begin{aligned}(ST)_k &= x_0 y_0 + (x_0 y_1 + x_1 y_0) + \dots + (x_0 y_k + \dots + x_k y_0)\\ &= x_0 T_k + x_1 T_{k-1} + \dots + x_k T_0\\ &= x_0(t_0 + \tau_k) + x_1(t_0 + \tau_{k-1}) + \dots + x_k(t_0 + \tau_0)\\ &= S_k t_0 + x_0 \tau_k + x_1 \tau_{k-1} + \dots + x_k \tau_0.\end{aligned}$$

Since $\lim_{k\to\infty} S_k t_0 = s_0 t_0$ by part (i), this part of the result will follow if we can show that

$$\lim_{k\to\infty}(x_0\tau_k + x_1\tau_{k-1} + \dots + x_k\tau_0) = 0. \tag{2.6}$$

Denote

$$\sigma = \sum_{j=0}^{\infty} |x_j|,$$

and for $\epsilon \in \mathbb{R}_{>0}$ choose $N_1 \in \mathbb{Z}_{>0}$ such that $|\tau_j| \leq \frac{\epsilon}{2\sigma}$ for $j \geq N_1$, this being possible since $(\tau_j)_{j\in\mathbb{Z}_{>0}}$ clearly converges to zero. Then, for $k \geq N_1$,

$$\begin{aligned}|x_0\tau_k + x_1\tau_{k-1} + \dots + x_k\tau_0| &\leq |x_0\tau_k + \dots + x_{k-N_1-1}\tau_{N_1-1}| + |x_{k-N_1}\tau_{N_1} + \dots + x_k\tau_0|\\ &\leq \tfrac{\epsilon}{2} + |x_{k-N_1}\tau_{N_1} + \dots + x_k\tau_0|.\end{aligned}$$

Since $\lim_{k\to\infty} x_k = 0$, choose $N_2 \in \mathbb{Z}_{>0}$ such that

$$|x_{k-N_1}\tau_{N_1} + \dots + x_k\tau_0| < \tfrac{\epsilon}{2}$$

for $k \geq N_2$. Then

$$\begin{aligned}\limsup_{k\to\infty} |x_0\tau_k + x_1\tau_{k-1} + \cdots + x_k\tau_0| &= \lim_{k\to\infty} \sup\{|x_0\tau_j + x_1\tau_{j-1} + \cdots + x_j\tau_0| \mid j \geq k\}\\ &\leq \lim_{k\to\infty} \sup\{\tfrac{\epsilon}{2} + |x_{k-N_1}\tau_{N_1} + \cdots + x_k\tau_0| \mid j \geq k\}\\ &\leq \sup\{\tfrac{\epsilon}{2} + |x_{k-N_1}\tau_{N_1} + \cdots + x_k\tau_0| \mid j \geq N_2\} \leq \epsilon.\end{aligned}$$

Thus

$$\limsup_{k\to\infty} |x_0\tau_k + x_1\tau_{k-1} + \cdots + x_k\tau_0| \leq 0,$$

and since clearly

$$\liminf_{k\to\infty} |x_0\tau_k + x_1\tau_{k-1} + \cdots + x_k\tau_0| \geq 0,$$

we infer that (2.6) holds by Proposition 2.3.17.

(vi) The reader can prove this as Exercise **??**. ■

The reader is recommended to remember the Cauchy product when we talk about convolution of discrete-time signals in Section 11.1.3.

missing stuff

2.4.7 Series with arbitrary index sets

It will be helpful on a few occasions to be able to sum series whose index set is not necessarily countable, and here we indicate how this can be done. This material should be considered optional until one comes to that point in the text where it is needed.

2.4.31 Definition (Sum of series for arbitrary index sets) Let A be a set and let $(x_a)_{a\in A}$ be a family of elements of $\overline{\mathbb{R}}$. Let $A_+ = \{a \in A \mid x_a \in [0,\infty]\}$ and $A_- = \{a \in A \mid x_a \in [-\infty, 0]\}$.

(i) If $x_a \in [0,\infty]$ for $a \in A$, then $\sum_{a\in A} x_a = \sup\{\sum_{a\in A'} x_a \mid A' \subseteq A \text{ is finite}\}$.

(ii) For a general family, $\sum_{a\in A} x_a = \sum_{a_+\in A_+} x_{a_+} - \sum_{a_-\in A_-}(-x_{a_-})$, provided that at least one of $\sum_{a_+\in A_+} x_{a_+}$ or $\sum_{a_-\in A_-}(-x_{a_-})$ is finite.

(iii) If both $\sum_{a_+\in A_+} x_{a_+}$ are $\sum_{a_-\in A_-}(-x_{a_-})$ are finite, then $(x_a)_{a\in A}$ is ***summable***. •

We should understand the relationship between this sort of summation and our existing notion of the sum of a series in the case where the index set is $\mathbb{Z}_{>0}$.

2.4.32 Proposition (A summable series with index set $\mathbb{Z}_{>0}$ is absolutely convergent) *A sequence $(\mathrm{x}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ is summable if and only if the series $\mathrm{S} = \sum_{j=1}^{\infty} \mathrm{x}_j$ is absolutely convergent.*

Proof Consider the sequences $(x_j^+)_{j\in\mathbb{Z}_{>0}}$ and $(x_j^-)_{j\in\mathbb{Z}_{>0}}$ defined by

$$x_j^+ = \max\{x_j, 0\}, \quad x_j^- = \max\{-x_j, 0\}, \qquad j \in \mathbb{Z}_{>0}.$$

Then $(x_j)_{j\in\mathbb{Z}_{>0}}$ is summable if and only if both of the expressions

$$\sup\Big\{\sum_{j\in A'} x_j^+ \;\Big|\; A' \subseteq \mathbb{Z}_{>0} \text{ is finite}\Big\}, \quad \sup\Big\{\sum_{j\in A'} x_j^- \;\Big|\; A' \subseteq \mathbb{Z}_{>0} \text{ is finite}\Big\} \tag{2.7}$$

are finite.

First suppose that $(x_j)_{j\in\mathbb{Z}_{>0}}$ is summable. Therefore, if $(S_k^+)_{k\in\mathbb{Z}_{>0}}$ and $(S_k^-)_{k\in\mathbb{Z}_{>0}}$ are the sequences of partial sums

$$S_k^+ = \sum_{j=1}^{k} x_j^+, \quad S_k^- = \sum_{j=1}^{k} x_j^-,$$

then these sequences are increasing and so convergent by (2.7). Then, by Proposition 2.3.23,

$$\sum_{j=1}^{\infty}|x_j| = \sum_{j=1}^{\infty} x_j^+ + \sum_{j=1}^{\infty} x_j^-$$

giving absolute convergence of S.

Now suppose that S is absolutely convergent. Then the subsets $\{S_k^+ \mid k\in\mathbb{Z}_{>0}\}$ and $\{S_k^- \mid k\in\mathbb{Z}_{>0}\}$ are bounded above (as well as being bounded below by zero) so that both expressions

$$\sup\{S_k^+ \mid k\in\mathbb{Z}_{>0}\}, \quad \sup\{S_k^- \mid k\in\mathbb{Z}_{>0}\}$$

are finite. Then for any finite set $A'\subseteq\mathbb{Z}_{>0}$ we have

$$\sum_{j\in A'} x_j^+ \le S^+_{\sup A'}, \quad \sum_{j\in A'} x_j^- \le S^-_{\sup A'}.$$

From this we deduce that

$$\sup\Big\{\sum_{j\in A'} x_j^+ \;\Big|\; A'\subseteq\mathbb{Z}_{>0} \text{ is finite}\Big\} \le \sup\{S_k^+ \mid k\in\mathbb{Z}_{>0}\},$$
$$\sup\Big\{\sum_{j\in A'} x_j^- \;\Big|\; A'\subseteq\mathbb{Z}_{>0} \text{ is finite}\Big\} \le \sup\{S_k^- \mid k\in\mathbb{Z}_{>0}\},$$

which implies that $(x_j)_{j\in\mathbb{Z}_{>0}}$ is summable. ■

Now we can actually show that, for a summable family of real numbers, only countably many of them can be nonzero.

2.4.33 Proposition (A summable family has at most countably many nonzero members) *If* $(\mathrm{x}_\mathrm{a})_{\mathrm{a}\in\mathrm{A}}$ *is summable, then the set* $\{\mathrm{a}\in\mathrm{A} \mid \mathrm{x}_\mathrm{a}\neq 0\}$ *is countable.*

Proof Note that for any $k\in\mathbb{Z}_{>0}$, the set $\{a\in A \mid |x_a|\ge\frac{1}{k}\}$ must be finite if $(x_a)_{a\in A}$ is summable (why?). Thus, since

$$\{a\in A \mid |x_a|\neq 0\} = \cup_{k\in\mathbb{Z}_{>0}}\{a\in A \mid |x_a|\ge\tfrac{1}{k}\},$$

the set $\{a\in A \mid x_a\neq 0\}$ is a countable union of finite sets, and so is countable by Proposition **??**. ■

A legitimate question is, since a summable family reduces to essentially being countable, why should we bother with the idea at all? The reason is simply that it will be notationally convenient in Section 3.3.4.

2.4.8 Notes

The numbers e and π are not only irrational, but have the much stronger property of being ***transcendental***. This means that they are not the roots of any polynomial having rational coefficients (see Definition **??**). That e is transcendental was proved by Hermite[7] in 1873, and the that π is transcendental was proved by Lindemann[8] in 1882.

The proof we give for the irrationality of π is essentially that of **IN:47**; this is the most commonly encountered proof, and is simpler than the original proof of Lambert[9] presented to the Berlin Academy in 1768.

Exercises

2.4.1 Let $S = \sum_{j=1}^{\infty} x_j$ be a series in $\mathbb{R}$, and, for $j \in \mathbb{Z}_{>0}$, define

$$x_j^+ = \max\{x_j, 0\}, \quad x_j^- = \max\{0, -x_j\}.$$

Show that, if S is conditionally convergent, then the series $S^+ = \sum_{j=1}^{\infty} x_j^+$ and $S^- = \sum_{j=1}^{\infty} x_j^-$ diverge to ∞.

2.4.2 In this exercise we consider more carefully the paradox of Zeno given in Exercise 1.5.2. Let us attach some symbols to the relevant data, so that we can say useful things. Suppose that the tortoise travels with constant velocity v_t and that Achilles travels with constant velocity v_a. Suppose that the tortoise gets a head start of t_0 seconds.

(a) Compute directly using elementary physics (i.e., time/distance/velocity relations) the time at which Achilles will overtake the tortoise, and the distance both will have travelled during that time.

(b) Consider the sequences $(d_j)_{j\in\mathbb{Z}_{>0}}$ and $(t_j)_{j\in\mathbb{Z}_{>0}}$ defined so that

1. d_1 is the distance travelled by the tortoise during the head start time t_0,
2. t_j, $j \in \mathbb{Z}_{>0}$, is the time it takes Achilles to cover the distance d_j,
3. d_j, $j \geq 2$, is the distance travelled by the tortoise in time t_{j-1}.

Find explicit expressions for these sequences in terms of t_0, v_t, and v_a.

(c) Show that the series $\sum_{j=1}^{\infty} d_j$ and $\sum_{j=1}^{\infty} t_j$ converge, and compute their limits.

(d) What is the relationship between the limits of the series in part (c) and the answers to part (a).

[7] Charles Hermite (1822–1901) was a French mathematician who made contributions to the fields of number theory, algebra, differential equations, and analysis.

[8] Carl Louis Ferdinand von Lindemann (1852–1939) was born in what is now Germany. His mathematical contributions were in the areas of analysis and geometry. He also was interested in physics.

[9] Johann Heinrich Lambert (1728–1777) was born in France. His mathematical work included contributions to analysis, geometry, and probability. He also made contributions to astronomical theory.

(e) Does this shed some light on how to resolve Zeno's paradox?

2.4.3 Show that

$$\Big|\sum_{j=1}^{m} x_j\Big| \leq \sum_{j=1}^{m} |x_j|$$

for any finite family $(x_1, \dots, x_m) \subseteq \mathbb{R}$.

2.4.4 State the correct version of Proposition 2.4.4 in the case that $S = \sum_{j=1}^{\infty} x_j$ is not absolutely convergent, and indicate why it is not a very interesting result.

2.4.5 For a sum

$$S = \sum_{j=1}^{\infty} s_j,$$

answer the following questions.

(a) Show that if S converges then the sequence $(s_j)_{j\in\mathbb{Z}_{>0}}$ converges to 0.

(b) Is the converse of part (a) true? That is to say, if the sequence $(s_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero, does S converge? If this is true, prove it. If it is not true, give a counterexample.

2.4.6 Do the following.

(a) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which converges in $\mathbb{R}$.

(b) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which diverges to ∞.

(c) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which diverges to $-\infty$.

(d) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}\left|\frac{x_{j+1}}{x_j}\right| = 1$ and which is oscillatory.

2.4.7 Do the following.

(a) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}|x_j|^{1/j} = 1$ and which converges in $\mathbb{R}$.

(b) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}|x_j|^{1/j} = 1$ and which diverges to ∞.

(c) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}|x_j|^{1/j} = 1$ and which diverges to $-\infty$.

(d) Find a series $\sum_{j=1}^{\infty} x_j$ for which $\lim_{j\to\infty}|x_j|^{1/j} = 1$ and which is oscillatory.

The next exercise introduces the notion of the decimal expansion of a real number. An ***infinite decimal expansion*** is a series in $\mathbb{Q}$ of the form

$$\sum_{j=0}^{\infty} \frac{a_j}{10^j}$$

where $a_0 \in \mathbb{Z}$ and where $a_j \in \{0, 1, \dots, 9\}$, $j \in \mathbb{Z}_{>0}$. An infinite decimal expansion is ***eventually periodic*** if there exists $k, m \in \mathbb{Z}_{>0}$ such that $a_{j+k} = a_j$ for all $j \geq m$.

2.4.8 (a) Show that the sequence of partial sums for an infinite decimal expansion is a Cauchy sequence.

(b) Show that, for every Cauchy sequence $(q_j)_{j\in\mathbb{Z}_{>0}}$, there exists a sequence $(d_j)_{j\in\mathbb{Z}_{>0}}$ of partial sums for a decimal expansion having the property that $[(q_j)_{j\in\mathbb{Z}_{>0}}] = [(d_j)_{j\in\mathbb{Z}_{>0}}]$ (the equivalence relation is that in the Cauchy sequences in $\mathbb{Q}$ as defined in Definition 2.1.16).

(c) Give an example that shows that two distinct infinite decimal expansions can be equivalent.

(d) Show that if two distinct infinite decimal expansions are equivalent, and if one of them is eventually periodic, then the other is also eventually periodic.

The previous exercises show that every real number is the limit of a (not necessarily unique) infinite decimal expansion. The next exercises characterise the infinite decimal expansions that correspond to rational numbers. First you will show that an eventually periodic decimal expansion corresponds to a rational number. Let $\sum_{j=0}^{\infty} \frac{a_j}{10^j}$ be an eventually periodic infinite decimal expansion and let $k, m \in \mathbb{Z}_{>0}$ have the property that $a_{j+k} = a_j$ for $j \geq m$. Denote by $x \in \mathbb{R}$ the number to which the infinite decimal expansion converges.

(e) Show that

$$10^{m+k}x = \sum_{j=0}^{\infty} \frac{b_j}{10^j}, \quad 10^m x = \sum_{j=0}^{\infty} \frac{c_j}{10^j}$$

are decimal expansions, and give expressions for b_j and c_j, $j \in \mathbb{Z}_{>0}$, in terms of a_j, $j \in \mathbb{Z}_{>0}$. In particular, show that $b_j = c_j$ for $j \geq 1$.

(f) Conclude that $(10^{m+k} - 10^m)x$ is an integer, and so x is therefore rational.

Next you will show that the infinite decimal expansion of a rational number is eventually periodic. Thus let $q \in \mathbb{Q}$.

(g) Let $q = \frac{a}{b}$ for $a, b \in \mathbb{Z}$ and with $b > 0$. For $j \in \{0, 1, \dots, b\}$, let $r_j \in \{0, 1, \dots, b-1\}$ satisfy $\frac{10^j}{b} = s_j + \frac{r_j}{b}$ for $s_j \in \mathbb{Z}$, i.e., r_j is the remainder after dividing 10^j by b. Show that at least two of the numbers $\{r_0, r_1, \dots, r_b\}$ must agree, i.e., conclude that $r_m = r_{m+k}$ for $k, m \in \mathbb{Z}_{\geq 0}$ satisfying $0 \leq m < m + k \leq b$.
Hint: *There are only* b *possible values for these* b + 1 *numbers.*

(h) Show that b exactly divides $10^{m+k} - 10^k$ with k and m as above. Thus $bc = 10^{m+k} - 10^k$ for some $c \in \mathbb{Z}$.

(i) Show that

$$\frac{a}{b} = 10^{-m} \frac{ac}{10^k - 1},$$

and so write

$$q = 10^{-m}\Big(s + \frac{r}{10^k - 1}\Big)$$

for $s \in \mathbb{Z}$ and $r \in \{0, 1, \dots, 10^k - 1\}$, i.e., r is the remainder after dividing ac by $10^k - 1$.

(j) Argue that we can write

$$b = \sum_{j=1}^{k} b_j 10^j,$$

for $b_j \in \{0, 1, \dots, 9\}$, $j \in \{1, \dots, k\}$.

(k) With b_j, $j \in \{1, \dots, k\}$ as above, define an infinite decimal expansion $\sum_{j=0}^{\infty} \frac{a_j}{10^j}$ by asking that $a_0 = 0$, that $a_j = b_j$, $j \in \{1, \dots, k\}$, and that $a_{j+km} = a_j$ for $j, m \in \mathbb{Z}_{>0}$. Let $d \in \mathbb{R}$ be the number to which this decimal expansion converges. Show that $(10^k - 1)d = b$, so $d \in \mathbb{Q}$.

(l) Show that $10^m q = s + d$, and so conclude that $10^m q$ has the eventually periodic infinite decimal expansion $s + \sum_{j=1}^{\infty} \frac{a_j}{10^j}$.

(m) Conclude that q has an eventually periodic infinite decimal expansion, and then conclude from (d) that any infinite decimal expansion for q is eventually periodic.

Section 2.5

Subsets of $\mathbb{R}$

In this section we study in some detail the nature of various sorts of subsets of $\mathbb{R}$. The character of these subsets will be of some importance when we consider the properties of functions defined on $\mathbb{R}$, and/or taking values in $\mathbb{R}$. Our presentation also gives us an opportunity to introduce, in a fairly simple setting, some concepts that will appear later in more abstract settings, e.g., open sets, closed sets, compactness.

Do I need to read this section? Unless you know the material here, it is indeed a good idea to read this section. Many of the ideas are basic, but some are not (e.g., the Heine–Borel Theorem). Moreover, many of the not-so-basic ideas will appear again later, particularly in Chapter **??**, and if a reader does not understand the ideas in the simple case of $\mathbb{R}$, things will only get more difficult. Also, the ideas expressed here will be essential in understanding even basic things about signals as presented in Chapter 8. •

2.5.1 Open sets, closed sets, and intervals

One of the basic building blocks in the understanding of the real numbers is the idea of an open set. In this section we define open sets and some related notions, and provide some simple properties associated to these ideas.

First, it is convenient to introduce the following ideas.

2.5.1 Definition (Open ball, closed ball) For $r \in \mathbb{R}_{>0}$ and $x_0 \in \mathbb{R}$,

(i) the ***open ball*** in $\mathbb{R}$ of radius r about x_0 is the set

$$\mathsf{B}(r, x_0) = \{x \in \mathbb{R} \mid |x - x_0| < r\},$$

and

(ii) the ***closed ball*** of radius r about x_0 is the set

$$\overline{\mathsf{B}}(r, x_0) = \{x \in \mathbb{R} \mid |x - x_0| \leq r\}.$$ •

These sets are simple to understand, and we depict them in Figure 2.3. With

Figure 2.3 An open ball (left) and a closed ball (right) in $\mathbb{R}$

the notion of an open ball, it is easy to give some preliminary definitions.

2.5.2 Definition (Open and closed sets in $\mathbb{R}$) A set $A \subseteq \mathbb{R}$ is:

(i) ***open*** if, for every $x \in A$, there exists $\epsilon \in \mathbb{R}_{>0}$ such that $\mathsf{B}(\epsilon, x) \subseteq A$ (the empty set is also open, by declaration);

(ii) ***closed*** if $\mathbb{R} \setminus A$ is open. •

A trivial piece of language associated with an open set is the notion of a neighbourhood.

2.5.3 Definition (Neighbourhood in $\mathbb{R}$) A ***neighbourhood*** of an element $x \in \mathbb{R}$ is an open set U for which $x \in U$. •

Some authors allow a "neighbourhood" to be a set A which contains a neighbourhood in our sense. Such authors will then frequently call what we call a neighbourhood an "open neighbourhood."

Let us give some examples of sets that are open, closed, or neither. The examples we consider here are important ones, since they are all examples of ***intervals***, which will be of interest at various times, and for various reasons, throughout these volumes. In particular, the notation we introduce here for intervals will be used a great deal.

2.5.4 Examples (Intervals)

1. For $a, b \in \mathbb{R}$ with $a < b$ the set
$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$$
is open. Indeed, let $x \in (a, b)$ and let $\epsilon = \frac{1}{2}\min\{b - x, x - a\}$. It is then easy to see that $\mathsf{B}(\epsilon, x) \subseteq (a, b)$. If $a \geq b$ we take the convention that $(a, b) = \emptyset$.

2. For $a \in \mathbb{R}$ the set
$$(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$$
is open. For example, if $x \in (a, \infty)$ then, if we define $\epsilon = \frac{1}{2}(x - a)$, we have $\mathsf{B}(\epsilon, x) \subseteq (a, \infty)$.

3. For $b \in \mathbb{R}$ the set
$$(-\infty, b) = \{x \in \mathbb{R} \mid x < b\}$$
is open.

4. For $a, b \in \mathbb{R}$ with $a \leq b$ the set
$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$$
is closed. Indeed, $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$. The sets $(-\infty, a)$ and (b, ∞) are both open, as we have already seen. Moreover, it is easy to see, directly from the definition, that the union of open sets is also an open set. Therefore, $\mathbb{R} \setminus [a, b]$ is open, and so $[a, b]$ is closed.

5. For $a \in \mathbb{R}$ the set
$$[a, \infty) = \{x \in \mathbb{R} \mid a \leq x\}$$
is closed since it complement in $\mathbb{R}$ is $(-\infty, a)$ which is open.

6. For $b \in \mathbb{R}$ the set
$$(-\infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$$
is closed.

7. For $a, b \in \mathbb{R}$ with $a < b$ the set
$$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$$
is neither open nor closed. To see that it is not open, note that $b \in (a, b]$, but that any open ball about b will contain points not in $(a, b]$. To see that $(a, b]$ is not closed, note that $a \in \mathbb{R} \setminus (a, b]$, and that any open ball about a will contain points not in $\mathbb{R} \setminus (a, b]$.

8. For $a, b \in \mathbb{R}$ with $a < b$ the set
$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$$
is neither open nor closed.

9. The set $\mathbb{R}$ is both open and closed. That it is open is clear. That it is closed follows since $\mathbb{R} \setminus \mathbb{R} = \emptyset$, and $\emptyset$ is, by convention, open. We will sometimes, although not often, write $\mathbb{R} = (-\infty, \infty)$. •

We shall frequently denote typical interval by I, and the set of intervals we denote by $\mathscr{I}$. If I and J are intervals with $J \subseteq I$, we will say that J is a ***subinterval*** of I. The expressions "open interval" and "closed interval" have their natural meanings as intervals that are, as subsets of $\mathbb{R}$, open and closed, respectively. An interval that is neither open nor closed will be called ***half-open*** or ***half-closed***. A ***left endpoint*** (resp. ***right endpoint***) for an interval I is a number $x \in \mathbb{R}$ such that $\inf I = x$ (resp. $\sup I = x$). An endpoint x, be it left or right, is ***open*** if $x \notin I$ and is ***closed*** if $x \in I$. If $\inf I = -\infty$ (resp. $\sup I = \infty$), then we saw that I is ***unbounded on the left*** (resp. ***unbounded on the right***). We will also use the interval notation to denote subsets of the extended real numbers $\overline{\mathbb{R}}$. Thus, we may write

1. $(a, \infty] = (a, \infty) \cup \{\infty\}$,
2. $[a, \infty] = [a, \infty) \cup \{\infty\}$,
3. $[-\infty, b) = (-\infty, b) \cup \{-\infty\}$,
4. $[-\infty, b] = (-\infty, b] \cup \{-\infty\}$, and
5. $[-\infty, \infty] = (-\infty, \infty) \cup \{-\infty, \infty\} = \overline{\mathbb{R}}$.

The following characterisation of intervals is useful.

2.5.5 Proposition (Characterisation of intervals) *A subset* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval if and only if, for each* $\mathrm{a}, \mathrm{b} \in \mathrm{I}$ *with* $\mathrm{a} < \mathrm{b}$, $[\mathrm{a}, \mathrm{b}] \subseteq \mathrm{I}$.

Proof It is clear from the definition that, if I is an interval, then, for each $a, b \in I$ with $a < b$, $[a, b] \subseteq I$. So suppose that, for each $a, b \in I$ with $a < b$, $[a, b] \subseteq I$. Let $A = \inf I$ and let $B = \sup I$. We have the following cases to consider.

1. $A = B$: Trivially I is an interval.

2. $A, B \in \mathbb{R}$ and $A \neq B$: Choose $a_1, b_1 \in I$ such that $a_1 < b_1$. Define $a_{j+1}, b_{j+1} \in I$, $j \in \mathbb{Z}_{>0}$, inductively as follows. Let a_{j+1} be a point in I to the left of $\frac{1}{2}(A + a_j)$ and let b_{j+1} be a point in I to the right of $\frac{1}{2}(b_j + B)$. These constructions make sense by definition of A and B. Note that $(a_j)_{j\in\mathbb{Z}_{>0}}$ is a monotonically decreasing sequence converging to A and that $(b_j)_{j\in\mathbb{Z}_{>0}}$ is a monotonically increasing sequence converging to B. Also,
$$\bigcup_{j\in\mathbb{Z}_{>0}} [a_j, b_j] \subseteq I.$$
We also have either $\cup_{j\in\mathbb{Z}_{>0}}[a_j, b_j] = (A, B)$, $\cup_{j\in\mathbb{Z}_{>0}}[a_j, b_j] = [A, B)$, $\cup_{j\in\mathbb{Z}_{>0}}[a_j, b_j] = (A, B]$, or $\cup_{j\in\mathbb{Z}_{>0}}[a_j, b_j] = [A, B]$. Therefore we conclude that I is an interval with endpoints A and B.
3. $A = -\infty$ and $B \in \mathbb{R}$. Choose $a_1, b_1 \in I$ with $a_a < b_1 < B$. Define $a_{j+1}, b_{j+1} \in I$, $j \in \mathbb{Z}_{>0}$, inductively by asking that a_{j+1} be a point in I to the left of $a_j - 1$ and that b_{j+1} be a point in I to the right of $\frac{1}{2}(b_j + B)$. These constructions make sense by definition of A and B. Thus $(a_j)_{j\in\mathbb{Z}_{>0}}$ is a monotonically decreasing sequence in I diverging to $-\infty$ and $(b_j)_{j\in\mathbb{Z}_{>0}}$ is a monotonically increasing sequence in I converging to B. Thus
$$\bigcup_{j\in\mathbb{Z}_{>0}} [a_j, b_j] =\subseteq I.$$
Note that either $\bigcup_{j\in\mathbb{Z}_{>0}}[a_j, b_j] = (-\infty, B)$ or $\bigcup_{j\in\mathbb{Z}_{>0}}[a_j, b_j] = (-\infty, B]$. This means that either $I = (-\infty, B)$ or $I = (-\infty, B]$.
4. $A \in \mathbb{R}$ and $B = \infty$: A construction entirely like the preceding one shows that either $I = (A, \infty)$ or $I = [A, \infty)$.
5. $A = -\infty$ and $B = \infty$: Choose $a_1, b_1 \in I$ with $a_1 < b_1$. Inductively define $a_{j+1}, b_{j+1} \in I$, $j \in \mathbb{Z}_{>0}$, by asking that a_{j+1} be a point in I to the left of a_j and that b_{j+1} be a point in I to the right of b_j. We then conclude that
$$\bigcup_{j\in\mathbb{Z}_{>0}} [a_j, b_j] = \mathbb{R} =\subseteq I,$$
and so $I = \mathbb{R}$.

In all cases we have concluded that I is an interval. ■

The following property of open sets will be useful for us, and tells us a little about the character of open sets.

2.5.6 Proposition (Open sets in $\mathbb{R}$ are unions of open intervals) *If* $\mathrm{U} \subseteq \mathbb{R}$ *is a nonempty open set then* U *is a countable union of disjoint open intervals.*

Proof Let $x \in U$ and let I_x be the largest open interval containing x and contained in U. This definition of I_x makes sense since the union of open intervals containing x is also an open interval containing x. Now to each interval can be associated a rational number within the interval. Therefore, the number of intervals to cover U can be associated with a subset of $\mathbb{Q}$, and is therefore countable or finite. This shows that U is indeed a finite or countable union of open intervals. ■

2.5.2 Partitions of intervals

In this section we consider the idea of partitioning an interval of the form $[a, b]$. This is a construction that will be useful in a variety of places, but since we dealt with intervals in the previous section, this is an appropriate time to make the definition and the associated constructions.

2.5.7 Definition (Partition of an interval) A ***partition*** of an interval $[a, b]$ is a family $(I_1, \dots, I_k)$ of intervals such that

(i) $\mathrm{int}(I_j) \neq \emptyset$ for $j \in \{1, \dots, k\}$,

(ii) $[a, b] = \cup_{j=1}^{k} I_j$, and

(iii) $I_j \cap I_l = \emptyset$ for $j \neq l$.

We denote by $\mathrm{Part}([a, b])$ the set of partitions of $[a, b]$. •

We shall always suppose that a partition $(I_1, \dots, I_k)$ is totally ordered so that the left endpoint of I_{j+1} agrees with the right endpoint of I_j for each $j \in \{1, \dots, k-1\}$. That is to say, when we write a partition, we shall list the elements of the set according to this total order. Note that associated to a partition $(I_1, \dots, I_k)$ are the endpoints of the intervals. Thus there exists a family $(x_0, x_1, \dots, x_k)$ of $[a, b]$, ordered with respect to the natural total order on $\mathbb{R}$, such that, for each $j \in \{1, \dots, k\}$, x_{j-1} is the left endpoint of I_j and x_j is the right endpoint of I_j. Note that necessarily we have $x_0 = a$ and $x_k = b$. The set of endpoints of the intervals in a partition $P = (I_1, \dots, I_k)$ we denote by $\mathrm{EP}(P)$. In Figure 2.4 we show a partition with all ingredients labelled. For a partition P with $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$, denote

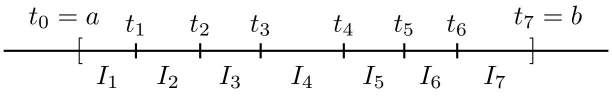


Figure 2.4 A partition

$$|P| = \max\{|x_j - x_l| \mid j, l \in \{1, \dots, k\}\},$$

which is the ***mesh*** of P. Thus $|P|$ is the length of the largest interval of the partition.

It is often useful to be able to say one partition is finer than another, and the following definition makes this precise.

2.5.8 Definition (Refinement of a partition) If P_1 and P_2 are partitions of an interval $[a, b]$, then P_2 is a ***refinement*** of P_1 if $\mathrm{EP}(P_1) \subseteq \mathrm{EP}(P_2)$. •

Next we turn to a sometimes useful construction involving the addition of certain structure onto a partition. This construction is rarely used in the text, so may be skipped until it is encountered.

2.5.9 Definition (Tagged partition, δ-fine tagged partition) Let $[a, b]$ be an interval and let $\delta\colon [a, b] \to \mathbb{R}_{>0}$.

(i) A ***tagged partition*** of $[a, b]$ is a finite family of pairs $((c_1, I_1), \dots, (c_k, I_k))$ where $(I_1, \dots, I_k)$ is a partition and where c_j is contained in the union of I_j with its endpoints.

(ii) A tagged partition $((c_1, I_1), \dots, (c_k, I_k))$ is ***δ-fine*** if the interval I_j, along with its endpoints, is a subset of $\mathsf{B}(\delta(c_j), c_j)$. •

The following result asserts that δ-fine tagged partitions always exist.

2.5.10 Proposition (δ-fine tagged partitions exist) *For any positive function $\delta\colon [\mathrm{a}, \mathrm{b}] \to \mathbb{R}_{>0}$, there exists a δ-fine tagged partition.*

Proof Let Δ be the set of all points $x \in (a, b]$ such that there exists a δ-fine tagged partition of $[a, x]$. Note that $(a, a + \delta(a)) \subseteq \Delta$ since, for each $x \in (a, a + \delta(a))$, $((a, [a, x]))$ is a δ-fine tagged partition of $[a, x]$. Let $b' = \sup \Delta$. We will show that $b' = b$ and that $b' \in \Delta$.

Since $b' = \sup \Delta$ there exists $b'' \in \Delta$ such that $b' - \delta(b') < b'' < b'$. Then there exists a δ-fine partition P' of $[a, b']$. Now $P' \cup ((b', (b'', b']))$ is δ-fine tagged partition of $[a, b']$. Thus $b' \in \Delta$.

Now suppose that $b' < b$ and choose $b'' < b$ such that $b' < b'' < b' + \delta(b')$. If P is a tagged partition of $[a, b']$ (this exists since $b' \in \Delta$), then $P \cup ((b', (b', b'']))$ is a δ-fine tagged partition of $[a, b'']$. This contradicts the fact that $b' = \sup \Delta$. Thus we conclude that $b' = b$. ■

2.5.3 Interior, closure, boundary, and related notions

Associated with the concepts of open and closed are a collection of useful concepts.

2.5.11 Definition (Accumulation point, cluster point, limit point in $\mathbb{R}$) Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is:

(i) an ***accumulation point*** for A if, for every neighbourhood U of x, the set $A \cap (U \setminus \{x\})$ is nonempty;

(ii) a ***cluster point*** for A if, for every neighbourhood U of x, the set $A \cap U$ is infinite;

(iii) a ***limit point*** of A if there exists a sequence $(x_j)_{j \in \mathbb{Z}_{>0}}$ in A converging to x.

The set of accumulation points of A is called the ***derived set*** of A, and is denoted by $\mathrm{der}(A)$. •

2.5.12 Remark (Conventions concerning “accumulation point,” “cluster point,” and “limit point”) There seems to be no agreed upon convention about what is meant by the three concepts of accumulation point, cluster point, and limit point. Some authors make no distinction between the three concepts at all. Some authors lump two together, but give the third a different meaning. As we shall see in Proposition 2.5.13 below, sometimes there is no need to distinguish between two of the concepts. However, in order to keep as clear as possible the transition to the more abstract presentation of Chapter **??**, we have gone with the most pedantic interpretation possible for the concepts of "accumulation point," "cluster point," and "limit point." •

The three concepts of accumulation point, cluster point, and limit point are actually excessive for $\mathbb{R}$ since, as the next result shall indicate, two of them are exactly the same. However, in the more general setup of Chapter **??**, the concepts are no longer equivalent.

2.5.13 Proposition ("Accumulation point" equals "cluster point" in $\mathbb{R}$) *For a set* $\mathrm{A} \subseteq \mathbb{R}$, $\mathrm{x} \in \mathbb{R}$ *is an accumulation point for* A *if and only if it is a cluster point for* A.

Proof It is clear that a cluster point for A is an accumulation point for A. Suppose that x is not a cluster point. Then there exists a neighbourhood U of x for which the set $A \cap U$ is finite. If $A \cap U = \{x\}$, then clearly x is not an accumulation point. If $A \cap U \neq \{x\}$, then $A \cap (U \setminus \{x\}) \supseteq \{x_1, \dots, x_k\}$ where the points $x_1, \dots, x_k$ are distinct from x. Now let

$$\epsilon = \tfrac{1}{2}\min\{|x_1 - x|, \dots, |x_k - x|\}.$$

Clearly $A \cap (\mathsf{B}(\epsilon, x) \setminus \{x\})$ is then empty, and so x is not an accumulation point for A. ■

Now let us give some examples that illustrate the differences between accumulation points (or equivalently cluster points) and limit points.

2.5.14 Examples (Accumulation points and limit points)

1. For any subset $A \subseteq \mathbb{R}$ and for every $x \in A$, x is a limit point for A. Indeed, the constant sequence $(x_j = x)_{j \in \mathbb{Z}_{>0}}$ is a sequence in A converging to x. However, as we shall see in the examples to follow, it is not the case that all points in A are accumulation points.
2. Let $A = (0, 1)$. The set of accumulation points of A is then easily seen to be $[0, 1]$. The set of limit points is also $[0, 1]$.
3. Let $A = [0, 1)$. Then, as in the preceding example, both the set of accumulation points and the set of limit points are the set $[0, 1]$.
4. Let $A = [0, 1] \cup \{2\}$. Then the set of accumulation points is $[0, 1]$ whereas the set of limit points is A.
5. Let $A = \mathbb{Q}$. One can readily check that the set of accumulation points of A is $\mathbb{R}$ and the set of limit points of A is also $\mathbb{R}$. •

The following result gives some properties of the derived set.

2.5.15 Proposition (Properties of the derived set in $\mathbb{R}$) *For* $\mathrm{A}, \mathrm{B} \subseteq \mathbb{R}$ *and for a family of subsets* $(\mathrm{A}_i)_{i \in I}$ *of* $\mathbb{R}$, *the following statements hold:*

(i) $\mathrm{der}(\emptyset) = \emptyset$;

(ii) $\mathrm{der}(\mathbb{R}) = \mathbb{R}$;

(iii) $\mathrm{der}(\mathrm{der}(\mathrm{A})) = \mathrm{der}(\mathrm{A})$;

(iv) *if* $\mathrm{A} \subseteq \mathrm{B}$ *then* $\mathrm{der}(\mathrm{A}) \subseteq \mathrm{der}(\mathrm{B})$;

(v) $\mathrm{der}(\mathrm{A} \cup \mathrm{B}) = \mathrm{der}(\mathrm{A}) \cup \mathrm{der}(\mathrm{B})$;

(vi) der(A ∩ B) ⊆ der(A) ∩ der(B).

Proof Parts (i) and (ii) follow directly from the definition of the derived set.

(iii) ***missing stuff***

(iv) Let $x \in \mathrm{der}(A)$ and let U be a neighbourhood of x. Then the set $A \cap (U \setminus \{x\})$ is nonempty, implying that the set $B \cap (U \setminus \{x\})$ is also nonempty. Thus $x \in \mathrm{der}(B)$.

(v) Let $x \in \mathrm{der}(A \cup B)$ and let U be a neighbourhood of x. Then the set $U \cap ((A \cup B) \setminus \{x\})$ is nonempty. But

$$\begin{aligned} U \cap ((A \cup B) \setminus \{x\}) &= U \cap ((A \setminus \{x\}) \cup (B \setminus \{x\})) \\ &= (U \cap (A \setminus \{x\})) \cup (U \cap (B \setminus \{x\})). \end{aligned} \tag{2.8}$$

Thus it cannot be that both $U \cap (A \setminus \{x\})$ and $U \cap (B \setminus \{x\})$ are empty. Thus x is an element of either $\mathrm{der}(A)$ or $\mathrm{der}(B)$.

Now let $x \in \mathrm{der}(A) \cup \mathrm{der}(A)$. Then, using (2.8), $U \cap ((A \cup B) \setminus \{x\})$ is nonempty, and so $x \in \mathrm{der}(A \cup B)$.

(vi) Let $x \in \mathrm{der}(A \cap B)$ and let U be a neighbourhood of x. Then $U \cap ((A \cap B) \setminus \{x\}) \neq \emptyset$. We have

$$U \cap ((A \cap B) \setminus \{x\}) = U \cap ((A \setminus \{x\}) \cap (B \setminus \{x\}))$$

Thus the sets $U \cap (A \setminus \{x\})$ and $U \cap (B \setminus \{x\})$ are both nonempty, showing that $x \in \mathrm{der}(A) \cap \mathrm{der}(B)$. ■

Next we turn to characterising distinguished subsets of subsets of $\mathbb{R}$.

2.5.16 Definition (Interior, closure, and boundary in $\mathbb{R}$) Let $A \subseteq \mathbb{R}$.

(i) The ***interior*** of A is the set

$$\mathrm{int}(A) = \cup\{U \mid U \subseteq A,\ U \text{ open}\}.$$

(ii) The ***closure*** of A is the set

$$\mathrm{cl}(A) = \cap\{C \mid A \subseteq C,\ C \text{ closed}\}.$$

(iii) The ***boundary*** of A is the set $\mathrm{bd}(A) = \mathrm{cl}(A) \cap \mathrm{cl}(\mathbb{R} \setminus A)$. •

In other words, the interior of A is the largest open set contained in A. Note that this definition makes sense since a union of open sets is open (Exercise 2.5.1). In like manner, the closure of A is the smallest closed set containing A, and this definition makes sense since an intersection of closed sets is closed (Exercise 2.5.1 again). Note that $\mathrm{int}(A)$ is open and $\mathrm{cl}(A)$ is closed. Moreover, since $\mathrm{bd}(A)$ is the intersection of two closed sets, it too is closed (Exercise 2.5.1 yet again).

Let us give some examples of interiors, closures, and boundaries.

2.5.17 Examples (Interior, closure, and boundary)

1. Let $A = \mathrm{int}(0,1)$. Then $\mathrm{int}(A) = (0,1)$ since A is open. We claim that $\mathrm{cl}(A) = [0,1]$. Clearly $[0,1] \subseteq \mathrm{cl}(A)$ since $[0,1]$ is closed and contains A. Moreover, the only smaller subsets contained in $[0,1]$ and containing A are $[0,1)$, $(0,1]$, and $(0,1)$, none of which are closed. We may then conclude that $\mathrm{cl}(A) = [0,1]$. Finally we claim that $\mathrm{bd}(A) = \{0,1\}$. To see this, note that we have $\mathrm{cl}(A) = [0,1]$ and $\mathrm{cl}(\mathbb{R} \setminus A) = (-\infty,0] \cup [1,\infty)$ (by an argument like that used to show that $\mathrm{cl}(A) = [0,1]$). Therefore, $\mathrm{bd}(A) = \mathrm{cl}(A) \cap \mathrm{cl}(\mathbb{R} \setminus A) = \{0,1\}$, as desired.

2. Let $A = [0,1]$. Then $\mathrm{int}(A) = (0,1)$. To see this, we note that $(0,1) \subseteq \mathrm{int}(A)$ since $(0,1)$ is open and contained in A. Moreover, the only larger sets contained in A are $[0,1)$, $(0,1]$, and $[0,1]$, none of which are open. Thus $\mathrm{int}(A) = (0,1)$, just as claimed. Since A is closed, $\mathrm{cl}(A) = A$. Finally we claim that $\mathrm{bd}(A) = \{0,1\}$. Indeed, $\mathrm{cl}(A) = [0,1]$ and $\mathrm{cl}(\mathbb{R} \setminus A) = (-\infty,0] \cup [1,\infty)$. Therefore, $\mathrm{bd}(A) = \mathrm{cl}(A) \cap \mathrm{cl}(\mathbb{R} \setminus A) = \{0,1\}$, as claimed.
3. Let $A = (0,1) \cup \{2\}$. We have $\mathrm{int}(A) = (0,1)$, $\mathrm{cl}(A) = [0,1] \cup \{2\}$, and $\mathrm{bd}(A) = \{0,1,2\}$. We leave the simple details of these assertions to the reader.
4. Let $A = \mathbb{Q}$. One readily ascertains that $\mathrm{int}(A) = \emptyset$, $\mathrm{cl}(A) = \mathbb{R}$, and $\mathrm{bd}(A) = \mathbb{R}$. •

Now let us give a characterisation of interior, closure, and boundary that are often useful in practice. Indeed, we shall often use these characterisations without explicitly mentioning that we are doing so.

2.5.18 Proposition (Characterisation of interior, closure, and boundary in $\mathbb{R}$) *For* $A \subseteq \mathbb{R}$, *the following statements hold:*

(i) $x \in \mathrm{int}(A)$ *if and only if there exists a neighbourhood* U *of* x *such that* $U \subseteq A$*;*

(ii) $x \in \mathrm{cl}(A)$ *if and only if, for each neighbourhood* U *of* x*, the set* $U \cap A$ *is nonempty;*

(iii) $x \in \mathrm{bd}(A)$ *if and only if, for each neighbourhood* U *of* x*, the sets* $U \cap A$ *and* $U \cap (\mathbb{R} \setminus A)$ *are nonempty.*

Proof (i) Suppose that $x \in \mathrm{int}(A)$. Since $\mathrm{int}(A)$ is open, there exists a neighbourhood U of x contained in $\mathrm{int}(A)$. Since $\mathrm{int}(A) \subseteq A$, $U \subseteq A$.

Next suppose that $x \notin \mathrm{int}(A)$. Then, by definition of interior, for any open set U for which $U \subseteq A$, $x \notin U$.

(ii) Suppose that there exists a neighbourhood U of x such that $U \cap A = \emptyset$. Then $\mathbb{R} \setminus U$ is a closed set containing A. Thus $\mathrm{cl}(A) \subseteq \mathbb{R} \setminus U$. Since $x \notin \mathbb{R} \setminus U$, it follows that $x \notin \mathrm{cl}(A)$.

Suppose that $x \notin \mathrm{cl}(A)$. Then x is an element of the open set $\mathbb{R} \setminus \mathrm{cl}(A)$. Thus there exists a neighbourhood U of x such that $U \subseteq \mathbb{R} \setminus \mathrm{cl}(A)$. In particular, $U \cap A = \emptyset$.

(iii) This follows directly from part (ii) and the definition of boundary. ■

Now let us state some useful properties of the interior of a set.

2.5.19 Proposition (Properties of interior in $\mathbb{R}$) *For* $A, B \subseteq \mathbb{R}$ *and for a family of subsets* $(A_i)_{i \in I}$ *of* $\mathbb{R}$*, the following statements hold:*

(i) $\mathrm{int}(\emptyset) = \emptyset$*;*

(ii) $\mathrm{int}(\mathbb{R}) = \mathbb{R}$*;*

(iii) $\mathrm{int}(\mathrm{int}(A)) = \mathrm{int}(A)$*;*

(iv) *if* $A \subseteq B$ *then* $\mathrm{int}(A) \subseteq \mathrm{int}(B)$*;*

(v) $\mathrm{int}(A \cup B) \supseteq \mathrm{int}(A) \cup \mathrm{int}(B)$*;*

(vi) $\mathrm{int}(A \cap B) = \mathrm{int}(A) \cap \mathrm{int}(B)$*;*

(vii) $\mathrm{int}(\cup_{i \in I} A_i) \supseteq \cup_{i \in I} \mathrm{int}(A_i)$*;*

(viii) $\mathrm{int}(\cap_{i \in I} A_i) \subseteq \cap_{i \in I} \mathrm{int}(A_i)$.

Moreover, a set $A \subseteq \mathbb{R}$ *is open if and only if* $\mathrm{int}(A) = A$.

Proof Parts (i) and (ii) are clear by definition of interior. Part (v) follows from part (vii), so we will only prove the latter.

(iii) This follows since the interior of an open set is the set itself.

(iv) Let $x \in \mathrm{int}(A)$. Then there exists a neighbourhood U of x such that $U \subseteq A$. Thus $U \subseteq B$, and the result follows from Proposition 2.5.18.

(vi) Let $x \in \mathrm{int}(A) \cap \mathrm{int}(B)$. Since $\mathrm{int}(A) \cap \mathrm{int}(B)$ is open by Exercise 2.5.1, there exists a neighbourhood U of x such that $U \subseteq \mathrm{int}(A) \cap \mathrm{int}(B)$. Thus $U \subseteq A \cap B$. This shows that $x \in \mathrm{int}(A \cap B)$. This part of the result follows from part (viii).

(vii) Let $x \in \cup_{i\in I} \mathrm{int}(A_i)$. By Exercise 2.5.1 the set $\cup_{i\in I} \mathrm{int}(A_i)$ is open. Thus there exists a neighbourhood U of x such that $U \subseteq \cup_{i\in I} \mathrm{int}(A_i)$. Thus $U \subseteq \cup_{i\in I} A_i$, from which we conclude that $x \in \mathrm{int}(\cup_{i\in I} A_i)$.

(viii) Let $x \in \mathrm{int}(\cap_{i\in I} A_i)$. Then there exists a neighbourhood U of x such that $U \subseteq \cap_{i\in I} A_i$. It therefore follows that $U \subseteq A_i$ for each $i \in I$, and so that $x \in \mathrm{int}(A_i)$ for each $i \in I$.

The final assertion follows directly from Proposition 2.5.18. ■

Next we give analogous results for the closure of a set.

2.5.20 Proposition (Properties of closure in $\mathbb{R}$) *For $A, B \subseteq \mathbb{R}$ and for a family of subsets $(A_i)_{i\in I}$ of $\mathbb{R}$, the following statements hold:*

(i) $\mathrm{cl}(\emptyset) = \emptyset$;

(ii) $\mathrm{cl}(\mathbb{R}) = \mathbb{R}$;

(iii) $\mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$;

(iv) *if* $A \subseteq B$ *then* $\mathrm{cl}(A) \subseteq \mathrm{cl}(B)$;

(v) $\mathrm{cl}(A \cup B) = \mathrm{cl}(A) \cup \mathrm{cl}(B)$;

(vi) $\mathrm{cl}(A \cap B) \subseteq \mathrm{cl}(A) \cap \mathrm{cl}(B)$;

(vii) $\mathrm{cl}(\cup_{i\in I} A_i) \supseteq \cup_{i\in I} \mathrm{cl}(A_i)$;

(viii) $\mathrm{cl}(\cap_{i\in I} A_i) \subseteq \cap_{i\in I} \mathrm{cl}(A_i)$.

Moreover, a set $A \subseteq \mathbb{R}$ is closed if and only if $\mathrm{cl}(A) = A$.

Proof Parts (i) and (ii) follow immediately from the definition of closure. Part (vi) follows from part (viii), so we will only prove the latter.

(iii) This follows since the closure of a closed set is the set itself.

(iv) Suppose that $x \in \mathrm{cl}(A)$. Then, for any neighbourhood U of x, the set $U \cap A$ is nonempty, by Proposition 2.5.18. Since $A \subseteq B$, it follows that $U \cap B$ is also nonempty, and so $x \in \mathrm{cl}(B)$.

(v) Let $x \in \mathrm{cl}(A \cup B)$. Then, for any neighbourhood U of x, the set $U \cap (A \cup B)$ is nonempty by Proposition 2.5.18. By Proposition 1.1.4, $U \cap (A \cup B) = (U \cap A) \cup (U \cap B)$. Thus the sets $U \cap A$ and $U \cap B$ are not both nonempty, and so $x \in \mathrm{cl}(A) \cup \mathrm{cl}(B)$. That $\mathrm{cl}(A) \cup \mathrm{cl}(B) \subseteq \mathrm{cl}(A \cup B)$ follows from part (vii).

(vi) Let $x \in \mathrm{cl}(A \cap B)$. Then, for any neighbourhood U of x, the set $U \cap (A \cap B)$ is nonempty. Thus the sets $U \cap A$ and $U \cap B$ are nonempty, and so $x \in \mathrm{cl}(A) \cap \mathrm{cl}(B)$.

(vii) Let $x \in \cup_{i\in I} \mathrm{cl}(A_i)$ and let U be a neighbourhood of x. Then, for each $i \in I$, $U \cap A_i \neq \emptyset$. Therefore, $\cup_{i\in I}(U \cap A_i) \neq \emptyset$. By Proposition 1.1.7, $\cup_{i\in I}(U \cap A_i) = U \cap (\cup_{i\in I} A_i)$, showing that $U \cap (\cup_{i\in I} A_i) \neq \emptyset$. Thus $x \in \mathrm{cl}(\cup_{i\in I} A_i)$.

(viii) Let $x \in \mathrm{cl}(\cap_{i\in I} A_i)$ and let U be a neighbourhood of x. Then the set $U \cap (\cap_{i\in I} A_i)$ is nonempty. This means that, for each $i \in I$, the set $U \cap A_i$ is nonempty. Thus $x \in \mathrm{cl}(A_i)$ for each $i \in I$, giving the result. ■

Note that there is a sort of "duality" between int and cl as concerns their interactions with union and intersection. This is reflective of the fact that open and closed sets themselves have such a "duality," as can be seen from Exercise 2.5.1. We refer the reader to Exercise 2.5.4 to construct counterexamples to any missing opposite inclusions in Propositions 2.5.19 and 2.5.20.

Let us state some relationships between certain of the concepts we have thus far introduced.

2.5.21 Proposition (Joint properties of interior, closure, boundary, and derived set in $\mathbb{R}$) *For* $A \subseteq \mathbb{R}$, *the following statements hold:*

(i) $\mathbb{R} \setminus \mathrm{int}(A) = \mathrm{cl}(\mathbb{R} \setminus A)$;

(ii) $\mathbb{R} \setminus \mathrm{cl}(A) = \mathrm{int}(\mathbb{R} \setminus A)$.

(iii) $\mathrm{cl}(A) = A \cup \mathrm{bd}(A)$;

(iv) $\mathrm{int}(A) = A - \mathrm{bd}(A)$;

(v) $\mathrm{cl}(A) = \mathrm{int}(A) \cup \mathrm{bd}(A)$;

(vi) $\mathrm{cl}(A) = A \cup \mathrm{der}(A)$;

(vii) $\mathbb{R} = \mathrm{int}(A) \cup \mathrm{bd}(A) \cup \mathrm{int}(\mathbb{R} \setminus A)$.

Proof (i) Let $x \in \mathbb{R} \setminus \mathrm{int}(A)$. Since $x \notin \mathrm{int}(A)$, for every neighbourhood U of x it holds that $U \not\subseteq A$. Thus, for any neighbourhood U of x, we have $U \cap (\mathbb{R} \setminus A) \neq \emptyset$, showing that $x \in \mathrm{cl}(\mathbb{R} \setminus A)$.

Now let $x \in \mathrm{cl}(\mathbb{R} \setminus A)$. Then for any neighbourhood U of x we have $U \cap (\mathbb{R} \setminus A) \neq \emptyset$. Thus $x \notin \mathrm{int}(A)$, so $x \in \mathbb{R} \setminus A$.

(ii) The proof here strongly resembles that for part (i), and we encourage the reader to provide the explicit arguments.

(iii) This follows from part (v).

(iv) Clearly $\mathrm{int}(A) \subseteq A$. Suppose that $x \in A \cap \mathrm{bd}(A)$. Then, for any neighbourhood U of x, the set $U \cap (\mathbb{R} \setminus A)$ is nonempty. Therefore, no neighbourhood of x is a subset of A, and so $x \notin \mathrm{int}(A)$. Conversely, if $x \in \mathrm{int}(A)$ then there is a neighbourhood U of x such that $U \subseteq A$. The precludes the set $U \cap (\mathbb{R} \setminus A)$ from being nonempty, and so we must have $x \notin \mathrm{bd}(A)$.

(v) Let $x \in \mathrm{cl}(A)$. For a neighbourhood U of x it then holds that $U \cap A \neq \emptyset$. If there exists a neighbourhood V of x such that $V \subseteq A$, then $x \in \mathrm{int}(A)$. If there exists *no* neighbourhood V of x such that $V \subseteq A$, then for every neighbourhood V of x we have $V \cap (\mathbb{R} \setminus A) \neq \emptyset$, and so $x \in \mathrm{bd}(A)$.

Now let $x \in \mathrm{int}(A) \cup \mathrm{bd}(A)$. If $x \in \mathrm{int}(A)$ then $x \in A$ and so $x \in\subseteq \mathrm{cl}(A)$. If $x \in \mathrm{bd}(A)$ then it follows immediately from Proposition 2.5.18 that $x \in \mathrm{cl}(A)$.

(vi) Let $x \in \mathrm{cl}(A)$. If $x \notin A$ then, for every neighbourhood U of x, $U \cap A = U \cap (A \setminus \{x\}) \neq \emptyset$, and so $x \in \mathrm{der}(A)$.

If $x \in A \cup \mathrm{der}(A)$ then either $x \in A \subseteq \mathrm{cl}(A)$, or $x \notin A$. In this latter case, $x \in \mathrm{der}(A)$ and so the set $U \cap (A \setminus \{x\})$ is nonempty for each neighbourhood U of x, and we again conclude that $x \in \mathrm{cl}(A)$.

(vii) Clearly $\operatorname{int}(A) \cap \operatorname{int}(\mathbb{R} \setminus A) = \emptyset$ since $A \cap (\mathbb{R} \setminus A) = \emptyset$. Now let $x \in \mathbb{R} \setminus (\operatorname{int}(A) \cup \operatorname{int}(\mathbb{R} \setminus A))$. Then, for any neighbourhood U of x, we have $U \not\subset A$ and $U \not\subset (\mathbb{R} \setminus A)$. Thus the sets $U \cap (\mathbb{R} \setminus A)$ and $U \cap A$ must both be nonempty, from which we conclude that $x \in \operatorname{bd}(A)$. ■

An interesting class of subset of $\mathbb{R}$ is the following.

2.5.22 Definition (Discrete subset of $\mathbb{R}$) A subset $A \subseteq \mathbb{R}$ is ***discrete*** if there exists $\epsilon \in \mathbb{R}_{>0}$ such that, for each $x, y \in A$, $|x - y| \geq \epsilon$. •

Let us give a characterisation of discrete sets.

2.5.23 Proposition (Characterisation of discrete sets in $\mathbb{R}$) *A discrete subset* A $\subseteq \mathbb{R}$ *is countable and has no accumulation points.*

Proof It is easy to show (Exercise 2.5.6) that if A is discrete and if $N \in \mathbb{Z}_{>0}$, then the set $A \cap [-N, N]$ is finite. Therefore

$$A = \cup_{N \in \mathbb{Z}_{>0}} A \cap [-N, N],$$

which gives A as a countable union of finite sets, implying that A is countable by Proposition **??**. Now let $\epsilon \in \mathbb{R}_{>0}$ satisfy $|x - y| \geq \epsilon$ for $x, y \in A$. Then, if $x \in A$ then the set $A \cap \mathsf{B}(\frac{\epsilon}{2}, x)$ is empty, implying that x is not an accumulation point. If $x \notin A$ then $\mathsf{B}(\frac{\epsilon}{2}, x)$ can contain at most one point from A, which again prohibits x from being an accumulation point. ■

The notion of a discrete set is actually a more general one having to do with what is known as the discrete topology (cf. Example **??–??**). The reader can explore some facts about discrete subsets of $\mathbb{R}$ in Exercise 2.5.6.

2.5.4 Compactness

The idea of compactness is absolutely fundamental in much of mathematics. The reasons for this are not at all clear to a newcomer to analysis. Indeed, the definition we give for compactness comes across as extremely unmotivated. This might be particularly since for $\mathbb{R}$ (or more generally, in $\mathbb{R}^n$) compact sets have a fairly banal characterisation as sets that are closed and bounded (Theorem 2.5.27). However, the original definition we give for a compact set is the most useful one. The main reason it is useful is that it allows for certain pointwise properties to be automatically extended to the entire set. A good example of this is Theorem 3.1.24, where continuity of a function on a compact set is extended to uniform continuity on the set. This idea of uniformity is an important one, and accounts for much of the value of the notion of compactness. But we are getting ahead of ourselves.

As indicated in the above paragraph, we shall give a rather strange seeming definition of compactness. Readers looking for a quick and dirty definition of compactness, valid for subsets of $\mathbb{R}$, can refer ahead to Theorem 2.5.27. Our construction relies on the following idea.

2.5.24 Definition (Open cover of a subset of $\mathbb{R}$) Let $A \subseteq \mathbb{R}$.

(i) An ***open cover*** for A is a family $(U_i)_{i\in I}$ of open subsets of $\mathbb{R}$ having the property that $A \subseteq \cup_{i\in I} U_i$.

(ii) A ***subcover*** of an open cover $(U_i)_{i\in I}$ of A is an open cover $(V_j)_{j\in J}$ of A having the property that $(V_j)_{j\in J} \subseteq (U_i)_{i\in I}$. •

The following property of open covers of subsets of $\mathbb{R}$ is useful.

2.5.25 Lemma (Lindelöf[10] Lemma for $\mathbb{R}$) *If* $(\mathrm{U_i})_{\mathrm{i}\in\mathrm{I}}$ *is an open cover of* $\mathrm{A} \subseteq \mathbb{R}$*, then there exists a countable subcover of* A.

Proof Let $\mathscr{B} = \{\mathsf{B}(r, x) \mid x, r \in \mathbb{Q}\}$. Note that $\mathscr{B}$ is a countable union of countable sets, and so is countable by Proposition **??**. Therefore, we can write $\mathscr{B} = (\mathsf{B}(r_j, x_j))_{j\in\mathbb{Z}_{>0}}$. Now define

$$\mathscr{B}' = \{\mathsf{B}(r_j, x_j) \mid \mathsf{B}(r_j, x_j) \subseteq U_i \text{ for some } i \in I\}.$$

Let us write $\mathscr{B}' = (\mathsf{B}(r_{j_k}, x_{j_k}))_{k\in\mathbb{Z}_{>0}}$. We claim that $\mathscr{B}'$ covers A. Indeed, if $x \in A$ then $x \in U_i$ for some $i \in I$. Since U_i is open there then exists $k \in \mathbb{Z}_{>0}$ such that $x \in \mathsf{B}(r_{j_k}, x_{j_k}) \subseteq U_i$. Now, for each $k \in \mathbb{Z}_{>0}$, let $i_k \in I$ satisfy $\mathsf{B}(r_{j_k}, x_{j_k}) \subseteq U_{i_k}$. Then the countable collection of open sets $(U_{i_k})_{k\in\mathbb{Z}_{>0}}$ clearly covers A since $\mathscr{B}'$ covers A. ■

Now we define the important notion of compactness, along with some other related useful concepts.

2.5.26 Definition (Bounded, compact, and totally bounded in $\mathbb{R}$) A subset $A \subseteq \mathbb{R}$ is:

(i) ***bounded*** if there exists $M \in \mathbb{R}_{>0}$ such that $A \subseteq \overline{\mathsf{B}}(M, 0)$;

(ii) ***compact*** if every open cover $(U_i)_{i\in I}$ of A possesses a finite subcover;

(iii) ***precompact***[11] if $\mathrm{cl}(A)$ is compact;

(iv) ***totally bounded*** if, for every $\epsilon \in \mathbb{R}_{>0}$ there exists $x_1, \dots, x_k \in \mathbb{R}$ such that $A \subseteq \cup_{j=1}^{k} \mathsf{B}(\epsilon, x_j)$. •

The simplest characterisation of compact subsets of $\mathbb{R}$ is the following. We shall freely interchange our use of the word compact between the definition given in Definition 2.5.26 and the conclusions of the following theorem.

2.5.27 Theorem (Heine–Borel[12] Theorem in $\mathbb{R}$) *A subset* $\mathrm{K} \subseteq \mathbb{R}$ *is compact if and only if it is closed and bounded.*

Proof Suppose that K is closed and bounded. We first consider the case when $K = [a, b]$. Let $\mathscr{O} = (U_i)_{i\in I}$ be an open cover for $[a, b]$ and let

$$S_{[a,b]} = \{x \in \mathbb{R} \mid x \leq b \text{ and } [a, x] \text{ has a finite subcover in } \mathscr{O}\}.$$

[10] Ernst Leonard Lindelöf (1870–1946) was a Finnish mathematician who worked in the areas of differential equations and complex analysis.

[11] What we call "precompact" is very often called "relatively compact." However, we shall use the term "relatively compact" for something different.

[12] Heinrich Eduard Heine (1821–1881) was a German mathematician who worked mainly with special functions. Félix Edouard Justin Emile Borel (1871–1956) was a French mathematician, and he worked mainly in the area of analysis.

Note that $S_{[a,b]} \neq \emptyset$ since $a \in S_{[a,b]}$. Let $c = \sup S_{[a,b]}$. We claim that $c = b$. Suppose that $c < b$. Since $c \in [a, b]$ there is some $\bar{i} \in I$ such that $c \in U_{\bar{i}}$. As $U_{\bar{i}}$ is open, there is some $\epsilon \in \mathbb{R}_{>0}$ sufficiently small that $\mathsf{B}(\epsilon, c) \subseteq U_{\bar{i}}$. By definition of c, there exists some $x \in (c - \epsilon, c)$ for which $x \in S_{[a,b]}$. By definition of $S_{[a,b]}$ there is a finite collection of open sets $U_{i_1}, \dots, U_{i_m}$ from $\mathscr{O}$ which cover $[a, x]$. Therefore, the finite collection $U_{i_1}, \dots, U_{i_m}, U_{\bar{i}}$ of open sets covers $[a, c+\epsilon)$. This then contradicts the fact that $c = \sup S_{[a,b]}$, so showing that $b = \sup S_{[a,b]}$. The result follows by definition of $S_{[a,b]}$.

Now suppose that K is a general closed and bounded set. Then $K \subseteq [a, b]$ for some suitable $a, b \in \mathbb{R}$. Suppose that $\mathscr{O} = (U_i)_{i\in I}$ is an open cover of K, and define a new open cover $\tilde{\mathscr{O}} = \mathscr{O} \cup (\mathbb{R} \setminus K)$. Note that $\cup_{i\in I} U_i \cup (\mathbb{R} \setminus K) = \mathbb{R}$ showing that $\tilde{\mathscr{O}}$ is an open cover for $\mathbb{R}$, and therefore also is an open cover for $[a, b]$. By the first part of the proof, there exists a finite subset of $\tilde{\mathscr{O}}$ which covers $[a, b]$, and therefore also covers K. We must show that this finite cover can be chosen so as not to include the set $\mathbb{R} \setminus K$ as this set is not necessarily in $\mathscr{O}$. However, if $[a, b]$ is covered by $U_{i_1}, \dots, U_{i_k}, \mathbb{R} \setminus K$, then one sees that K is covered by $U_{i_1}, \dots, U_{i_k}$, since $K \cap (\mathbb{R} \setminus K) = \emptyset$. Thus we have arrived at a finite subset of $\mathscr{O}$ covering K, as desired.

Now suppose that K is compact. Consider the following collection of open subsets: $\mathscr{O}_K = (\mathsf{B}(\epsilon, x))_{x\in K}$. Clearly this is an open cover of K. Thus there exists a finite collection of point $x_1, \dots, x_k \in K$ such that $(\mathsf{B}(\epsilon, x_j))_{j\in\{1,\dots,k\}}$ covers K. If we take

$$M = \max\{|x_1|, \dots, |x_k|\} + 2$$

then we easily see that $K \subseteq \overline{\mathsf{B}}(M, 0)$, so that K is bounded. Now suppose that K is not closed. Then $K \subset \mathrm{cl}(K)$. By part (vi) of Proposition 2.5.21 there exists an accumulation point x_0 of K that is not in K. Then, for any $j \in \mathbb{Z}_{>0}$ there exists a point $x_j \in K$ such that $|x_0 - x_j| < \frac{1}{j}$. Define

$$U_j = (-\infty, x_0 - \tfrac{1}{j}) \cup (x_0 + \tfrac{1}{j}, \infty),$$

noting that U_j is open, since it is the union of open sets (see Exercise 2.5.1). We claim that $(U_j)_{j\in\mathbb{Z}_{>0}}$ is an open cover of K. Indeed, we will show that $\cup_{j\in\mathbb{Z}_{>0}} U_j = \mathbb{R} \setminus \{x_0\}$. To see this, let $x \in \mathbb{R} \setminus \{x_0\}$ and choose $k \in \mathbb{Z}_{>0}$ such that $\frac{1}{k} < |x - x_0|$. Then it follows by definition of U_k that $x \in U_k$. Since $x_0 \notin K$, we then have $K \subseteq \cup_{j\in\mathbb{Z}_{>0}} U_j$. Next we show that there is no finite subset of $(U_j)_{j\in\mathbb{Z}_{>0}}$ that covers K. Indeed, consider a finite set $j_1, \dots, j_k \in \mathbb{Z}_{>0}$, and suppose without loss of generality that $j_1 < \dots < j_k$. Then the point x_{j_k+1} satisfies $|x_0 - x_{j_k+1}| < \frac{1}{j_k+1} < \frac{1}{j_k}$, implying that $x_{j_k+1} \notin U_{j_k} \supseteq \dots \supseteq U_{j_1}$. Thus, if K is not closed, we have constructed an open cover of K having no finite subcover. From this we conclude that if K is compact, then it is closed. ■

The Heine–Borel Theorem has the following useful corollary.

2.5.28 Corollary (Closed subsets of compact sets in $\mathbb{R}$ are compact) *If* $\mathrm{A} \subseteq \mathbb{R}$ *is compact and if* $\mathrm{B} \subseteq \mathrm{A}$ *is closed, then* B *is compact.*

Proof Since A is bounded by the Heine–Borel Theorem, B is also bounded. Thus B is also compact, again by the Heine–Borel Theorem. ■

In Chapter **??** we shall encounter many of the ideas in this section in the more general setting of topological spaces. Many of the ideas for $\mathbb{R}$ transfer directly to this more general setting. However, with compactness, some care must be exercised. In particular, it is *not* true that, in a general topological space, a subset is compact if and only if it is closed and bounded. Indeed, in a general topological space, the notion of bounded is not defined. It is not an uncommon error for newcomers to confuse "compact" with "closed and bounded" in situations where this is not the case.

missing stuff

The following result is another equivalent characterisation of compact subsets of $\mathbb{R}$, and is often useful.

2.5.29 Theorem (Bolzano–Weierstrass[13] Theorem in $\mathbb{R}$) *A subset* $\mathrm{K} \subseteq \mathbb{R}$ *is compact if and only if every sequence in* K *has a subsequence which converges in* K*.*

Proof First suppose that K is compact. Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in K. Since K is bounded by Theorem 2.5.27, the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ is bounded. We next show that there exists either a monotonically increasing, or a monotonically decreasing, subsequence of $(x_j)_{j\in\mathbb{Z}_{>0}}$. Define

$$D = \{j \in \mathbb{Z}_{>0} \mid x_k > x_j,\ k > j\}$$

If the set D is infinite, then we can write $D = (j_k)_{k\in\mathbb{Z}_{>0}}$. By definition of D, it follows that $x_{j_{k+1}} > x_{j_k}$ for each $k \in \mathbb{Z}_{>0}$. Thus the subsequence $(x_{j_k})_{k\in\mathbb{Z}_{>0}}$ is monotonically increasing. If the set D is finite choose $j_1 > \sup D$. Then there exists $j_2 > j_1$ such that $x_{j_2} \leq x_{j_1}$. Since $j_2 > \sup D$, there then exists $j_3 > j_2$ such that $x_{j_3} \leq x_{j_2}$. By definition of D, this process can be repeated inductively to yield a monotonically decreasing subsequence $(x_{j_k})_{k\in\mathbb{Z}_{>0}}$. It now follows from Theorem 2.3.8 that the sequence $(x_{j_k})_{k\in\mathbb{Z}_{>0}}$, be it monotonically increasing or monotonically decreasing, converges.

Next suppose that every sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ in K possesses a convergent subsequence. Let $(U_i)_{i\in I}$ be an open cover of K, and by Lemma 2.5.25 choose a countable subcover which we denote by $(U_j)_{j\in\mathbb{Z}_{>0}}$. Now suppose that every finite subcover of $(U_j)_{j\in\mathbb{Z}_{>0}}$ does not cover K. This means that, for every $k \in \mathbb{Z}_{>0}$, the set $C_k = K \setminus \left(\cup_{j=1}^{k} U_j\right)$ is nonempty. Thus we may define a sequence $(x_k)_{k\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ such that $x_k \in C_k$. Since the sequence $(x_k)_{k\in\mathbb{Z}_{>0}}$ is in K, it possesses a convergent subsequence $(x_{k_m})_{m\in\mathbb{Z}_{>0}}$, by hypotheses. Let x be the limit of this subsequence. Since $x \in K$ and since $K = \cup_{j\in\mathbb{Z}_{>0}} U_j$, $x \in U_l$ for some $l \in \mathbb{Z}_{>0}$. Since the sequence $(x_{k_m})_{m\in\mathbb{Z}_{>0}}$ converges to x, it follows that there exists $N \in \mathbb{Z}_{>0}$ such that $x_{k_m} \in U_l$ for $m \geq N$. But this contradicts the definition of the sequence $(x_k)_{k\in\mathbb{Z}_{>0}}$, forcing us to conclude that our assumption is wrong that there is no finite subcover of K from the collection $(U_j)_{j\in\mathbb{Z}_{>0}}$. ■

The following property of compact intervals of $\mathbb{R}$ is useful.

2.5.30 Theorem (Lebesgue[14] number for compact intervals) *Let* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *be a compact interval. Then for any open cover* $(\mathrm{U}_\alpha)_{\alpha\in\mathrm{A}}$ *of* $[\mathrm{a}, \mathrm{b}]$*, there exists* $\delta \in \mathbb{R}_{>0}$*, called the*

[13] Bernard Placidus Johann Nepomuk Bolzano (1781–1848) was a Czechoslovakian philosopher, mathematician, and theologian who made mathematical contributions to the field of analysis. Karl Theodor Wilhelm Weierstrass (1815–1897) is one of the greatest of all mathematicians. He made significant contributions to the fields of analysis, complex function theory, and the calculus of variations.

[14] Henri Léon Lebesgue (1875–1941) was a French mathematician. His work was in the area of analysis. The Lebesgue integral is considered to be one of the most significant contributions to mathematics in the past century or so.

Lebesgue number *of* I*, such that, for each* $\mathrm{x} \in [\mathrm{a}, \mathrm{b}]$*, there exists* $\alpha \in \mathrm{A}$ *such that* $\mathsf{B}(\delta, \mathrm{x}) \cap \mathrm{I} \subseteq \mathrm{U}_\alpha$*.*

Proof Suppose there exists an open cover $(U_\alpha)_{\alpha \in A}$ such that, for all $\delta \in \mathbb{R}_{>0}$, there exists $x \in [a, b]$ such that none of the sets U_α, $\alpha \in A$, contains $\mathsf{B}(\delta, x) \cap I$. Then there exists a sequence $(x_j)_{j \in \mathbb{Z}_{>0}}$ in I such that

$$\{\alpha \in A \mid \mathsf{B}(\tfrac{1}{j}, x_j) \subseteq U_\alpha\} = \emptyset$$

for each $j \in \mathbb{Z}_{>0}$. By the Bolzano–Weierstrass Theorem there exists a subsequence $(x_{j_k})_{k \in \mathbb{Z}_{>0}}$ that converges to a point, say x, in $[a, b]$. Then there exists $\epsilon \in \mathbb{R}_{>0}$ and $\alpha \in A$ such that $\mathsf{B}(\epsilon, x) \subseteq U_\alpha$. Now let $N \in \mathbb{Z}_{>0}$ be sufficiently large that $|x_{j_k} - x| < \frac{\epsilon}{2}$ for $k \geq N$ and such that $\frac{1}{j_N} < \frac{\epsilon}{2}$. Now let $k \geq N$. Then, if $y \in \mathsf{B}(\frac{1}{j_k}, x_{j_k})$ we have

$$|y - x| = |y - x_{j_k} + x_{j_k} - x| \leq |y - x_{j_k}| + |x - x_{j_k}| < \epsilon.$$

Thus we arrive at the contradiction that $\mathsf{B}(\frac{1}{j_k}, x_{j_k}) \subseteq U_\alpha$. ■

The following result is sometimes useful.

2.5.31 Proposition (Countable intersections of nested compact sets are nonempty) *Let* $(\mathrm{K}_\mathrm{j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *be a collection of compact subsets of* $\mathbb{R}$ *satisfying* $\mathrm{K}_{\mathrm{j}+1} \subseteq \mathrm{K}_\mathrm{j}$*. Then* $\cap_{\mathrm{j} \in \mathbb{Z}_{>0}} \mathrm{K}_\mathrm{j}$ *is nonempty.*

Proof It is clear that $K = \cap_{j \in \mathbb{Z}_{>0}} K_j$ is bounded, and moreover it is closed by Exercise 2.5.1. Thus K is compact by the Heine–Borel Theorem. Let $(x_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence for which $x_j \in K_j$ for $j \in \mathbb{Z}_{>0}$. This sequence is thus a sequence in K_1 and so, by the Bolzano–Weierstrass Theorem, has a subsequence $(x_{j_k})_{k \in \mathbb{Z}_{>0}}$ converging to $x \in K_1$. The sequence $(x_{j_{k+1}})_{k \in \mathbb{Z}_{>0}}$ is then a sequence in K_2 which is convergent, so showing that $x \in K_2$. Similarly, one shows that $x \in K_j$ for all $j \in \mathbb{Z}_{>0}$, giving the result. ■

Finally, let us indicate the relationship between the notions of relative compactness and total boundedness. We see that for $\mathbb{R}$ these concepts are the same. This may not be true in general.***missing stuff***

2.5.32 Proposition ("Precompact" equals "totally bounded" in $\mathbb{R}$) *A subset of* $\mathbb{R}$ *is precompact if and only if it is totally bounded.*

Proof Let $A \subseteq \mathbb{R}$.

First suppose that A is precompact. Since $A \subseteq \mathrm{cl}(A)$ and since $\mathrm{cl}(A)$ is bounded by the Heine–Borel Theorem, it follows that A is bounded. It is then easy to see that A is totally bounded.

Now suppose that A is totally bounded. For $\epsilon \in \mathbb{R}_{>0}$ let $x_1, \dots, x_k \in \mathbb{R}$ have the property that $A \subseteq \cup_{j=1}^k \mathsf{B}(\epsilon, x_j)$. If

$$M_0 = \max\{|x_j - x_l| \mid j, l \in \{i, \dots, k\}\} + 2\epsilon,$$

then it is easy to see that $A \subseteq \mathsf{B}(M, 0)$ for any $M > M_0$. Then $\mathrm{cl}(A) \subseteq \overline{\mathsf{B}}(M, 0)$ by part (iv) of Proposition 2.5.20, and so $\mathrm{cl}(A)$ is bounded. Since $\mathrm{cl}(A)$ is closed, it follows from the Heine–Borel Theorem that A is precompact. ■

missing stuff missing stuff

2.5.5 Connectedness

The idea of a connected set will come up occasionally in these volumes. Intuitively, a set is connected if it cannot be "broken in two." We will study it more systematically in ***missing stuff***, and here we only give enough detail to effectively characterise connected subsets of $\mathbb{R}$.

2.5.33 Definition (Connected subset of $\mathbb{R}$) Subsets $A, B \subseteq \mathbb{R}$ are ***separated*** if $A \cap \mathrm{cl}(B) = \emptyset$ and $\mathrm{cl}(A) \cap B = \emptyset$. A subset $S \subseteq \mathbb{R}$ is ***disconnected*** if $S = A \cup B$ for nonempty separated subsets A and B. A subset $S \subseteq \mathbb{R}$ is ***connected*** if it is not disconnected. •

Rather than give examples, let us simply immediately characterise the connected subsets of $\mathbb{R}$, since this renders all examples trivial to understand.

2.5.34 Theorem (Connected subsets of $\mathbb{R}$ are intervals and vice versa) *A subset* S $\subseteq \mathbb{R}$ *is connected if and only if* S *is an interval.*

Proof Suppose that S is not an interval. Then, by Proposition 2.5.5, there exists $a, b \in S$ with $a < b$ and $c \in (a, b)$ such that $c \notin S$. Let $A_c = S \cap (-\infty, c)$ and $B_c = S \cap (c, \infty)$, and note that both A_c and B_c are nonempty. Also, since $c \notin S$, $S = A_c \cup B_c$. Since $(-\infty, c) \cap [c, \infty) = \emptyset$ and $(-\infty, c] \cap (c, \infty) = \emptyset$, A_c and B_c are separated. That S is not connected follows.

Now suppose that S is not connected, and write $S = A \cup B$ for nonempty separated sets A and B. Without loss of generality, let $a \in A$ and $b \in B$ have the property that $a < b$. Note that $A \cap [a, b]$ is bounded so that $c = \sup A \cap [a, b]$ exists in $\mathbb{R}$. Then $c \in \mathrm{cl}(A \cap [a, b]) \subseteq \mathrm{cl}(A) \cap [a, b]$. In other words, $c \in \mathrm{cl}(A)$. Since $\mathrm{cl}(A) \cap B = \emptyset$, $c \notin B$. If $c \notin A$ then $c \notin S$, and so S is not connected by Proposition 2.5.5. If $c \in A$ then, since $A \cap \mathrm{cl}(B) = \emptyset$, $c \notin \mathrm{cl}(B)$. In this case there exists an open interval containing c that does not intersect $\mathrm{cl}(B)$. In particular, there exists $d > c$ such that $d \notin B$. Since $d > c$ we also have $d \notin A$, and so $d \notin S$. Again we conclude that S is not an interval by Proposition 2.5.5. ■

Let us consider a few examples.

2.5.35 Examples (Connected subsets of sets)

1. If $D \subseteq \mathbb{R}$ is a discrete set as given in Definition 2.5.22. From Theorem 2.5.34 we see that the only subsets of D that are connected are singletons.
2. Note that it also follows from Theorem 2.5.34 that the only connected subsets of $\mathbb{Q} \subseteq \mathbb{R}$ are singletons. However, $\mathbb{Q}$ is not discrete. •

2.5.6 Sets of measure zero

The topic of this section will receive a full treatment in the context of measure theory as presented in Chapter 5. However, it is convenient here to talk about a simple concepts from measure theory, one which formalises the idea of a set being "small." We shall only give here the definition and a few examples. The reader should look ahead to Chapter 5 for more detail.

2.5.36 Definition (Set of measure zero in $\mathbb{R}$) A subset $A \subseteq \mathbb{R}$ has ***measure zero***, or is ***of measure zero***, if

$$\inf\Big\{\sum_{j=1}^{\infty}|b_j - a_j| \;\Big|\; A \subseteq \bigcup_{j\in\mathbb{Z}_{>0}}(a_j, b_j)\Big\} = 0. \qquad \bullet$$

The idea, then, is that one can cover a set A with open intervals, each of which have some length. One can add all of these lengths to get a total length for the intervals used to cover A. Now, if one can make this total length arbitrarily small, then the set has measure zero.

2.5.37 Notation ("Almost everywhere" and "a.e.") We give here an important piece of notation associated to the notion of a set of measure zero. Let $A \subseteq \mathbb{R}$ and let $P\colon A \to \{\mathrm{true}, \mathrm{false}\}$ be a property defined on A (see the prelude to the Principle of Transfinite Induction, Theorem **??**). The property P holds ***almost everywhere, a.e.***, or ***for almost every* x ∈ A** if the set $\{x \in A \mid P(x) = \mathrm{false}\}$ has measure zero. •

This is best illustrated with some examples.

2.5.38 Examples (Sets of measure zero)

1. Let $A = \{x_1, \dots, x_k\}$ for some distinct $x_1, \dots, x_k \in \mathbb{R}$. We claim that this set has measure zero. Note that for any $\epsilon \in \mathbb{R}_{>0}$ the intervals $(x_j - \frac{\epsilon}{4k}, x_j + \frac{\epsilon}{4k})$, $j \in \{1, \dots, k\}$, clearly cover A. Now consider the countable collection of open intervals

$$((x_j - \tfrac{\epsilon}{4k}, x_j + \tfrac{\epsilon}{4k}))_{j\in\{1,\dots,k\}} \cup ((0, \tfrac{\epsilon}{2^{j+1}}))_{j\in\mathbb{Z}_{>0}}$$

obtained by adding to the intervals covering A a collection of intervals around zero. The total length of these intervals is

$$\sum_{j=1}^{k}|(x_j + \tfrac{\epsilon}{4k}) - (x_j - \tfrac{\epsilon}{4k})| + \frac{\epsilon}{2}\sum_{j=1}^{\infty}\frac{1}{2^j} = \frac{\epsilon}{2} + \frac{\epsilon}{2},$$

using the fact that $\sum_{j=1}^{\infty}\frac{\epsilon}{2^j} = 1$ (by Example 2.4.2–1). Since $\inf\{2k\epsilon \mid \epsilon \in \mathbb{R}_{>0}\} = 0$, our claim that A has zero measure is validated.

2. Now let $A = \mathbb{Q}$ be the set of rational numbers. To show that A has measure zero, note that from Exercise 2.1.3 that A is countable. Thus we can write the elements of A as $(q_j)_{j\in\mathbb{Z}_{>0}}$. Now let $\epsilon \in \mathbb{R}_{>0}$ and for $j \in \mathbb{Z}_{>0}$ define $a_j = q_j - \frac{\epsilon}{2^j}$ and $b_j = q_j + \frac{\epsilon}{2^j}$. Then the collection (a_j, b_j), $j \in \mathbb{Z}_{>0}$, covers A. Moreover,

$$\sum_{j=1}^{\infty}|b_j - a_j| = \sum_{j=1}^{\infty}\frac{2\epsilon}{2^j} = 2\epsilon,$$

using the fact, shown in Example 2.4.2–1, that the series $\sum_{j=1}^{\infty}\frac{1}{2^j}$ converges to 1. Now, since $\inf\{2\epsilon \mid \epsilon \in \mathbb{R}_{>0}\} = 0$, it follows that A indeed has measure zero.

3. Let $A = \mathbb{R} \setminus \mathbb{Q}$ be the set of irrational numbers. We claim that this set does not have measure zero. To see this, let $k \in \mathbb{Z}_{>0}$ and consider the set $A_k = A \cap [-k, k]$. Now let $\epsilon \in \mathbb{R}_{>0}$. We claim that if $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$, is a collection of open intervals for which $A_k \subseteq \cup_{j\in\mathbb{Z}_{>0}}(a_j, b_j)$, then
$$\sum_{j=1}^{\infty} |b_j - a_j| \geq 2k - \epsilon. \tag{2.9}$$
To see this, let $((c_l, d_l))_{l\in\mathbb{Z}_{>0}}$ be a collection of intervals such that $\mathbb{Q} \cap [-k, k] \subseteq \cup_{l\in\mathbb{Z}_{>0}}(c_l, d_l)$ and such that
$$\sum_{l=1}^{\infty} |d_l - c_l| < \epsilon.$$
Such a collection of intervals exists since we have already shown that $\mathbb{Q}$, and therefore $\mathbb{Q} \cap [-k, k]$, has measure zero (see Exercise 2.5.7). Now note that
$$[-k, k] \subseteq \Big(\bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j)\Big) \cup \Big(\bigcup_{l\in\mathbb{Z}_{>0}} (c_l, d_l)\Big),$$
so that
$$\Big(\sum_{j=1}^{\infty} |b_j - a_j|\Big) + \Big(\sum_{l=1}^{\infty} |d_l - c_l|\Big) \geq 2k.$$
From this we immediately conclude that (2.9) does indeed hold. Moreover, (2.9) holds for every $k \in \mathbb{Z}_{>0}$, for every $\epsilon \in \mathbb{R}_{>0}$, and for every open cover $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ of A_k. Thus,
$$\begin{aligned} \inf\Big\{\sum_{l=1}^{\infty} |\tilde{b}_l - \tilde{a}_l| \;\Big|\; A \subseteq \bigcup_{l\in\mathbb{Z}_{>0}} (\tilde{a}_l, \tilde{b}_l)\Big\} \\ \geq \inf\Big\{\sum_{j=1}^{\infty} |b_j - a_j| \;\Big|\; A_k \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j)\Big\} \geq 2k - \epsilon \end{aligned}$$
for every $k \in \mathbb{Z}_{>0}$ and for every $\epsilon \in \mathbb{R}_{>0}$. This precludes A from having measure zero. •

The preceding examples suggest sets of measure zero are countable. This is not so, and the next famous example gives an example of an uncountable set with measure zero.

2.5.39 Example (An uncountable set of measure zero: the middle-thirds Cantor set) In this example we construct one of the standard "strange" sets used in real analysis to exhibit some of the characteristics that can possibly be attributed to subsets of ℝ. We shall also use this set in a construction in Example 3.2.27 to give an example of a continuous monotonically increasing function whose derivative is zero almost everywhere.

Let $C_0 = [0,1]$. Then define

$$\begin{aligned} C_1 &= [0,\tfrac{1}{3}] \cup [\tfrac{2}{3},1], \\ C_2 &= [0,\tfrac{1}{9}] \cup [\tfrac{2}{9},\tfrac{1}{3}] \cup [\tfrac{2}{3},\tfrac{7}{9}] \cup [\tfrac{8}{9},1], \\ &\vdots \end{aligned}$$

so that C_k is a collection of 2^k disjoint closed intervals each of length 3^{-k} (see Figure 2.5). We define $C = \cap_{k\in\mathbb{Z}_{>0}} C_k$, which we call the ***middle-thirds Cantor set***.

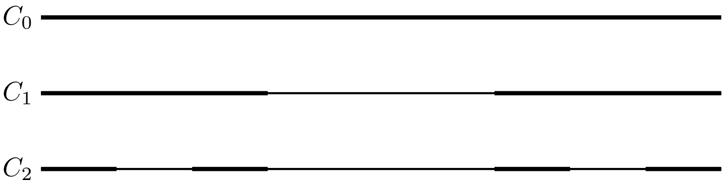


Figure 2.5 The first few sets used in the construction of the middle-thirds Cantor set

Let us give some of the properties of C.

1 **Lemma** C *has the same cardinality as* $[0,1]$.

Proof Note that each of the sets C_k, $k \in \mathbb{Z}_{\geq 0}$, is a collection of disjoint closed intervals. Let us write $C_k = \cup_{j=1}^{2^k} I_{k,j}$, supposing that the intervals $I_{k,j}$ are enumerated such that the right endpoint of $I_{k,j}$ lies to the left of the left endpoint of $I_{k,j+1}$ for each $k \in \mathbb{Z}_{\geq 0}$ and $j \in \{1,\dots,2^k\}$. Now note that each interval $I_{k+1,j}$, $k \in \mathbb{Z}_{\geq 0}$, $j \in \{1,\dots,2^{k+1}\}$ comes from assigning two intervals to each of the intervals $I_{k,j}$, $k \in \mathbb{Z}_{\geq 0}$, $j \in \{1,\dots,2^k\}$. Assign to an interval $I_{k+1,j}$, $k \in \mathbb{Z}_{\geq 0}$, $j \in \{1,\dots,2^k\}$, the number 0 (resp. 1) if it the left (resp. right) interval coming from an interval $I_{k,j'}$ of C_k. In this way, each interval in C_k, $k \in \mathbb{Z}_{\geq 0}$, is assigned a 0 or a 1 in a unique manner. Since, for each point in $x \in C$, there is exactly one $j \in \{1,\dots,2^k\}$ such that $x \in I_{k,j}$. Therefore, for each point in C there is a unique decimal expansion $0.n_1n_2n_3\dots$ where $n_k \in \{0,1\}$. Moreover, for every such decimal expansion, there is a corresponding point in C. However, such decimal expansions are exactly binary decimal expansions for points in $[0,1]$. In other words, there is a bijection from C to $[0,1]$. ▼

2 **Lemma** C *is a set of measure zero.*

Proof Let $\epsilon \in \mathbb{R}_{>0}$. Note that each of the sets C_k can be covered by a finite number of closed intervals whose lengths sum to $\left(\frac{2}{3}\right)^k$. Therefore, each of the sets C_k can be covered by open intervals whose lengths sum to $\left(\frac{2}{3}\right)^k + \frac{\epsilon}{2}$. Choosing k sufficiently large that $\left(\frac{2}{3}\right)^k < \frac{\epsilon}{2}$ we see that C is contained in the union of a finite collection of open intervals whose lengths sum to ϵ. Since ϵ is arbitrary, it follows that C has measure zero. ▼

This example thus shows that sets of measure zero, while "small" in some sense, can be "large" in terms of the number of elements they possess. Indeed, in terms of cardinality, C has the same size as $[0,1]$, although their measures differ by as much as possible. •

2.5.7 Cantor sets

The remainder of this section is devoted to a characterisation of certain sorts of exotic sets, perhaps the simplest example of which is the middle-thirds Cantor set of Example 2.5.39. This material is only used occasionally, and so can be omitted until the reader feels they need/want to understand it.

The qualifier "middle-thirds" in Example 2.5.39 makes one believe that there might be a general notion of a "Cantor set." This is indeed the case.

2.5.40 Definition (Cantor set) Let $I \subseteq \mathbb{R}$ be a closed interval. A subset $A \subseteq I$ is a ***Cantor set*** if

(i) A is closed,

(ii) $\operatorname{int}(A) = \emptyset$, and

(iii) every point of A is an accumulation point of A. •

We leave it to the reader to verify in Exercise 2.5.10 that the middle-thirds Cantor set is a Cantor set, according to the previous definition.

One might wonder whether all Cantor sets have the properties of having the cardinality of an interval and of having measure zero. To address this, we give a result and an example. The result shows that all Cantor sets are uncountable.

2.5.41 Proposition (Cantor sets are uncountable) *If* A $\subseteq \mathbb{R}$ *is a nonempty set having the property that each of its points is an accumulation point, then* A *is uncountable. In particular, Cantor sets are uncountable.*

Proof Any finite set has no accumulation points by Proposition 2.5.13. Therefore A must be either countably infinite or uncountable. Suppose that A is countable and write $A = (x_j)_{j\in\mathbb{Z}_{>0}}$. Let $y_1 \in A \setminus \{x_1\}$. For $r_1 < |x_1 - y_1|$ we have $x_1 \notin \overline{\mathsf{B}}(r_1, y_1)$. We note that y_1 is an accumulation point for $A \setminus \{x_1, x_2\}$; this follows immediately from Proposition 2.5.13. Thus there exists $y_2 \in A \setminus \{x_1, x_2\}$ such that $y_2 \in \mathsf{B}(r_1, y_1)$ and such that $y_2 \neq y_1$. If $r_2 < \min\{|x_2 - y_2|, r_1 - |y_2 - y_2|\}$ then $x_2 \notin \overline{\mathsf{B}}(r_2, y_2)$ and $\overline{\mathsf{B}}(r_2, y_2) \subseteq \mathsf{B}(r_1, y_1)$ by a simple application of the triangle inequality. Continuing in this way we define a sequence $(\overline{\mathsf{B}}(r_j, y_j))_{j\in\mathbb{Z}_{>0}}$ of closed balls having the following properties:

1. $\overline{\mathsf{B}}(r_{j+1}, y_{j+1}) \subseteq \overline{\mathsf{B}}(r_j, y_j)$ for each $j \in \mathbb{Z}_{>0}$;
2. $x_j \notin \overline{\mathsf{B}}(r_j, y_j)$ for each $j \in \mathbb{Z}_{>0}$.

Note that $(\overline{\mathsf{B}}(r_j, y_j) \cap A)_{j\in\mathbb{Z}_{>0}}$ is a nested sequence of compact subsets of A, and so by Proposition 2.5.31, $\cap_{j\in\mathbb{Z}_{>0}}(\overline{\mathsf{B}}(r_j, y_j) \cap A)$ is a nonempty subset of A. However, for any $j \in \mathbb{Z}_{>0}$, $x_j \notin \cap_{j\in\mathbb{Z}_{>0}}(\overline{\mathsf{B}}(r_j, y_j) \cap A)$, and so we arrive, by contradiction, to the conclusion that A is not countable. ■

The following example shows that Cantor sets may not have measure zero.

2.5.42 Example (A Cantor set not having zero measure) We will define a subset of $[0,1]$ that is a Cantor set, but does not have measure zero. The construction mirrors closely that of Example 2.5.39.

We let $\epsilon \in (0,1)$. Let $C_{\epsilon,0} = [0,1]$ and define $C_{\epsilon,1}$ by deleting from $C_{\epsilon,0}$ an open interval of length $\frac{\epsilon}{2}$ centered at the midpoint of $C_{\epsilon,0}$. Note that $C_{\epsilon,1}$ consists of two disjoint closed intervals whose lengths sum to $1 - \frac{\epsilon}{2}$. Next define $C_{\epsilon,2}$ by deleting from $C_{\epsilon,1}$ two open intervals, each of length $\frac{\epsilon}{8}$, centered at the midpoints of each of the intervals comprising $C_{\epsilon,1}$. Note that $C_{\epsilon,2}$ consists of four disjoint closed intervals whose lengths sum to $1 - \frac{\epsilon}{4}$. Proceed in this way, defining a sequence of sets $(C_{\epsilon,k})_{k\in\mathbb{Z}_{>0}}$, where $C_{\epsilon,k}$ consists of 2^k disjoint closed intervals whose lengths sum to $1 - \sum_{j=1}^{k} \frac{\epsilon}{2^j} = 1 - \epsilon$. Take $C_\epsilon = \cap_{k\in\mathbb{Z}_{>0}} C_{\epsilon,k}$.

Let us give the properties of C_ϵ in a series of lemmata.

1 Lemma *C_ϵ is a Cantor set.*

Proof That C_ϵ is closed follows from Exercise 2.5.1 and the fact that it is the intersection of a collection of closed sets. To see that $\mathrm{int}(C_\epsilon) = \emptyset$, let $I \subseteq [0,1]$ be an open interval and suppose that $I \subseteq C_\epsilon$. This means that $I \subseteq C_{\epsilon,k}$ for each $k \in \mathbb{Z}_{>0}$. Note that the sets $C_{\epsilon,k}$, $k \in \mathbb{Z}_{>0}$, are unions of closed intervals, and that for any $\delta \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that the lengths of the intervals comprising $C_{\epsilon,k}$ are less than δ for $k \geq N$. Thus the length of I must be zero, and so $I = \emptyset$. Thus C_ϵ contains no nonempty open intervals, and so must have an empty interior. To see that every point of C_ϵ is an accumulation point of C_ϵ, we note that all points in C_ϵ are endpoints for one of the closed intervals comprising $C_{\epsilon,k}$ for some $k \in \mathbb{Z}_{>0}$. Moreover, it is clear that every neighbourhood of a point in C_ϵ must contain another endpoint from one of the closed intervals comprising $C_{\epsilon,k}$ for some $k \in \mathbb{Z}_{>0}$. Indeed, were this not the case, this would imply the existence of a nonempty open interval contained in C_ϵ, and we have seen that there can be no such interval. ▼

2 Lemma *C_ϵ is uncountable.*

Proof This can be proved in exactly the same manner as the middle-thirds Cantor set was shown to be uncountable. ▼

3 Lemma *C_ϵ does not have measure zero.*

Proof Once one knows the basic properties of Lebesgue measure, it follows immediately that C_ϵ has, in fact, measure $1 - \epsilon$. However, since we have not yet defined measure, let us prove that C_ϵ does not have measure zero, using only the definition of a set of measure zero. Let $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ be a countable collection of open intervals having the property that

$$C_\epsilon \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j).$$

Since C_ϵ is closed, it is compact by Corollary 2.5.28. Therefore, there exists a finite collection $((a_{j_l}, b_{j_l}))_{l\in\{1,\dots,m\}}$ of intervals having the property that

$$C_\epsilon \subseteq \bigcup_{l=1}^{m} (a_{j_l}, b_{j_l}). \tag{2.10}$$

We claim that there exists $k \in \mathbb{Z}_{>0}$ such that

$$C_{\epsilon,k} \subseteq \bigcup_{l=1}^{m} (a_{j_l}, b_{j_l}). \tag{2.11}$$

Indeed, suppose that, for each $k \in \mathbb{Z}_{>0}$ there exists $x_k \in C_{\epsilon,k}$ such that $x_k \notin \cup_{l=1}^{m}(a_{j_l}, b_{j_l})$. The sequence $(x_k)_{k\in\mathbb{Z}_{>0}}$ is then a sequence in the compact set $C_{\epsilon,1}$, and so by the Bolzano–Weierstrass Theorem, possesses a subsequence $(x_{k_r})_{r\in\mathbb{Z}_{>0}}$ converging to $x \in C_{\epsilon,1}$. But the sequence $(x_{k_{r+1}})_{r\in\mathbb{Z}_{>0}}$ is then a convergent sequence in $C_{\epsilon,2}$, so $x \in C_{\epsilon,2}$. Continuing in this way, $x \in \cap_{k\in\mathbb{Z}_{>0}} C_{\epsilon,k}$. Moreover, the sequence $(x_k)_{k\in\mathbb{Z}_{>0}}$ is also a sequence in the closed set $[0,1] - \cup_{l=1}^{m}(a_{j_l}, b_{j_l})$, and so we conclude that $x \in [0,1] - \cup_{l=1}^{m}(a_{j_l}, b_{j_l})$. Thus we contradict the condition (2.10), and so there indeed must be a $k \in \mathbb{Z}_{>0}$ such that (2.11) holds. However, this implies that any collection of open intervals covering C_ϵ must have lengths which sum to at least $1 - \epsilon$. Thus C_ϵ cannot have measure zero. ▼

Cantor sets such as C_ϵ are sometimes called ***fat Cantor sets***, reflecting the fact that they do not have measure zero. Note, however, that they are not *that* fat, since they have an empty interior! •

2.5.8 Notes

Some uses of δ-fine tagged partitions in real analysis can be found in the paper of **RAG:98**.

Exercises

2.5.1 For an arbitrary collection $(U_a)_{a\in A}$ of open sets and an arbitrary collection $(C_b)_{b\in B}$ of closed sets, do the following:

(a) show that $\cup_{a\in A} U_a$ is open;

(b) show that $\cap_{b\in B} C_b$ is closed;

For open sets U_1 and U_2 and closed sets C_1 and C_2, do the following:

(c) show that $U_1 \cap U_2$ is open;

(d) show that $C_1 \cup C_2$ is closed.

2.5.2 Show that a set $A \subseteq \mathbb{R}$ is closed if and only if it contains all of its limit points.

2.5.3 For $A \subseteq \mathbb{R}$, show that $\mathrm{bd}(A) = \mathrm{bd}(\mathbb{R} \setminus A)$.

2.5.4 Find counterexamples to the following statements (cf. Propositions 2.5.15, 2.5.19, and 2.5.20):

(a) $\mathrm{int}(A \cup B) \subseteq \mathrm{int}(A) \cup \mathrm{int}(B)$;

(b) $\mathrm{int}(\cup_{i\in I} A_i) \subseteq \cup_{i\in I} \mathrm{int}(A_i)$;

(c) $\mathrm{int}(\cap_{i\in I} A_i) \supseteq \cap_{i\in I} \mathrm{int}(A_i)$;

(d) $\mathrm{cl}(A \cap B) \supseteq \mathrm{cl}(A) \cap \mathrm{cl}(B)$;

(e) $\mathrm{cl}(\cup_{i\in I} A_i) \subseteq \cup_{i\in I} \mathrm{cl}(A_i)$;

(f) $\mathrm{cl}(\cap_{i\in I} A_i) \supseteq \cap_{i\in I} \mathrm{cl}(A_i)$.

Hint: *No fancy sets are required. Intervals will suffice in all cases.*

2.5.5 For each of the following statements, prove the statement if it is true, and give a counterexample if it is not:

(a) $\mathrm{int}(A_1 \cup A_2) = \mathrm{int}(A_1) \cup \mathrm{int}(A_2)$;

(b) $\mathrm{int}(A_1 \cap A_2) = \mathrm{int}(A_1) \cap \mathrm{int}(A_2)$;

(c) $\mathrm{cl}(A_1 \cup A_2) = \mathrm{cl}(A_1) \cup \mathrm{cl}(A_2)$;

(d) $\mathrm{cl}(A_1 \cap A_2) = \mathrm{cl}(A_1) \cap \mathrm{cl}(A_2)$;

(e) $\mathrm{bd}(A_1 \cup A_2) = \mathrm{bd}(A_1) \cup \mathrm{bd}(A_2)$;

(f) $\mathrm{bd}(A_1 \cap A_2) = \mathrm{bd}(A_1) \cap \mathrm{bd}(A_2)$.

2.5.6 Do the following:

(a) show that any finite subset of $\mathbb{R}$ is discrete;

(b) show that a discrete bounded set is finite;

(c) find a set $A \subseteq \mathbb{R}$ that is countable and has no accumulation points, but that is not discrete.

2.5.7 Show that if $A \subseteq \mathbb{R}$ has measure zero and if $B \subseteq A$, then B has measure zero.

2.5.8 Show that any countable subset of $\mathbb{R}$ has measure zero.

2.5.9 Let $(Z_j)_{j\in\mathbb{Z}_{>0}}$ be a family of subsets of $\mathbb{R}$ that each have measure zero. Show that $\cup_{j\in\mathbb{Z}_{>0}} Z_j$ also has measure zero.

2.5.10 Show that the set C constructed in Example 2.5.39 is a Cantor set.

Chapter 3

Functions of a real variable

In the preceding chapter we endowed the set $\mathbb{R}$ with a great deal of structure. In this chapter we employ this structure to endow functions whose domain and range is $\mathbb{R}$ with some useful properties. These properties include the usual notions of continuity and differentiability given in first-year courses on calculus. The theory of the Riemann integral is also covered here, and it can be expected that students will have at least a functional familiarity with this. However, students who have had the standard engineering course (at least in North American universities) dealing with these topics will find the treatment here a little different than what they are used to. Moreover, there are also topics covered that are simply not part of the standard undergraduate curriculum, but which still fit under the umbrella of "functions of a real variable." These include a detailed discussion of functions of bounded variation, an introductory treatment of absolutely continuous functions, and a generalisation of the Riemann integral called the Riemann–Stieltjes integral.

Do I need to read this chapter? For readers having had a good course in analysis, this chapter can easily be bypassed completely. It can be expected that all other readers will have some familiarity with the material in this chapter, although not perhaps with the level of mathematical rigour we undertake. This level of mathematical rigour is not necessarily needed, if all one wishes to do is deal with $\mathbb{R}$-valued functions defined on $\mathbb{R}$ (as is done in most engineering undergraduate programs). However, we will wish to use the ideas introduced in this chapter, particularly those from Section 3.1, in contexts far more general than the simple one of $\mathbb{R}$-valued functions. Therefore, it will be helpful, at least, to understand the simple material in this chapter in the rigorous manner in which it is presented.

As for the more advanced material, such as is contained in Sections 3.3, **??**, and **??**, it is probably best left aside on a first reading. The reader will be warned when this material is needed in the presentation.

Some of what we cover in this chapter, particularly notions of continuity, differentiability, and Riemann integrability, will be covered in more generality in Chapter **??**. Aggressive readers may want to skip this material here and proceed directly to the more general case. •

Contents

3.1 Continuous $\mathbb{R}$-valued functions on $\mathbb{R}$ 137
3.1.1 Definition and properties of continuous functions 137
3.1.2 Discontinuous functions 141
3.1.3 Continuity and operations on functions 145
3.1.4 Continuity, and compactness and connectedness 147
3.1.5 Monotonic functions and continuity 150
3.1.6 Convex functions and continuity 153
3.1.7 Piecewise continuous functions 159
3.2 Differentiable $\mathbb{R}$-valued functions on $\mathbb{R}$ 163
3.2.1 Definition of the derivative 163
3.2.2 The derivative and continuity 167
3.2.3 The derivative and operations on functions 170
3.2.4 The derivative and function behaviour 175
3.2.5 Monotonic functions and differentiability 183
3.2.6 Convex functions and differentiability 190
3.2.7 Piecewise differentiable functions 196
3.2.8 Notes 197
3.3 $\mathbb{R}$-valued functions of bounded variation 199
3.3.1 Functions of bounded variation on compact intervals 199
3.3.2 Functions of bounded variation on general intervals 203
3.3.3 Bounded variation and operations on functions 206
3.3.4 Saltus functions 210
3.3.5 The saltus function for a function of locally bounded variation 213
3.4 The Riemann integral 217
3.4.1 Step functions 217
3.4.2 The Riemann integral on compact intervals 219
3.4.3 Characterisations of Riemann integrable functions on compact intervals 221
3.4.4 The Riemann integral on noncompact intervals 228
3.4.5 The Riemann integral and operations on functions 234
3.4.6 The Fundamental Theorem of Calculus and the Mean Value Theorems . 239
3.4.7 The Cauchy principal value 245
3.4.8 Notes 247
3.5 Sequences and series of $\mathbb{R}$-valued functions 248
3.5.1 Pointwise convergent sequences 248
3.5.2 Uniformly convergent sequences 249
3.5.3 Dominated and bounded convergent sequences 252
3.5.4 Series of $\mathbb{R}$-valued functions 254
3.5.5 Some results on uniform convergence of series 255
3.5.6 The Weierstrass Approximation Theorem 257
3.5.7 Swapping limits with other operations 263
3.5.8 Notes 266
3.6 Some $\mathbb{R}$-valued functions of interest 267
3.6.1 The exponential function 267
3.6.2 The natural logarithmic function 269
3.6.3 Power functions and general logarithmic functions 271

3.6.4 Trigonometric functions 275
3.6.5 Hyperbolic trigonometric functions 282

Section 3.1

Continuous $\mathbb{R}$-valued functions on $\mathbb{R}$

The notion of continuity is one of the most important in all of mathematics. Here we present this important idea in its simplest form: continuity for functions whose domain and range are subsets of $\mathbb{R}$.

Do I need to read this section? Unless you are familiar with this material, it is probably a good idea to read this section fairly carefully. It builds on the structure of $\mathbb{R}$ built up in Chapter 2 and uses this structure in an essential way. It is essential to understand this if one is to understand the more general ideas of continuity that will arise in Chapter **??**. This section also provides an opportunity to improve one's facility with the $\epsilon - \delta$ formalism. •

3.1.1 Definition and properties of continuous functions

In this section we will deal with functions defined on an interval $I \subseteq \mathbb{R}$. This interval might be open, closed, or neither, and bounded, unbounded, or neither. In this section, we shall reserve the letter I to denote such a general interval. It will also be convenient to say that a subset $A \subseteq I$ is ***open*** if $A = U \cap I$ for an open subset U of $\mathbb{R}$.[1] For example, if $I = [0, 1]$, then the subset $[0, \frac{1}{2})$ is an open subset of I, but not an open subset of $\mathbb{R}$. We will be careful to explicitly say that a subset is open *in* I if this is what we mean. *There is a chance for confusion here, so the reader is advised to be alert!*

Let us give the standard definition of continuity.

3.1.1 Definition (Continuous function) Let $I \subseteq \mathbb{R}$ be an interval. A map $f\colon I \to \mathbb{R}$ is:

(i) ***continuous at*** $\mathbf{x_0} \in \mathbf{I}$ if, for every $\epsilon \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ such that $|f(x) - f(x_0)| < \epsilon$ whenever $x \in I$ satisfies $|x - x_0| < \delta$;

(ii) ***continuous*** if it is continuous at each $x_0 \in I$;

(iii) ***discontinuous at*** $\mathbf{x_0} \in \mathbf{I}$ if it is not continuous at x_0;

(iv) ***discontinuous*** if it is not continuous. •

The idea behind the definition of continuity is this: one can make the values of a continuous function as close as desired by making the points at which the function is evaluated sufficiently close. Readers not familiar with the definition should be prepared to spend some time embracing it. An often encountered oversimplification of continuity is illustrated in Figure 3.1. The idea is supposed to be that the function whose graph is shown on the left is continuous because its graph has no "gaps," whereas the function on the right is discontinuous because its graph does have a "gap." As we shall see in Example 3.1.2–4 below, it is

[1] This is entirely related to the notion of relative topology which we will discuss in Section **??** for sets of multiple real variables and in Definition **??** within the general context of topological spaces.

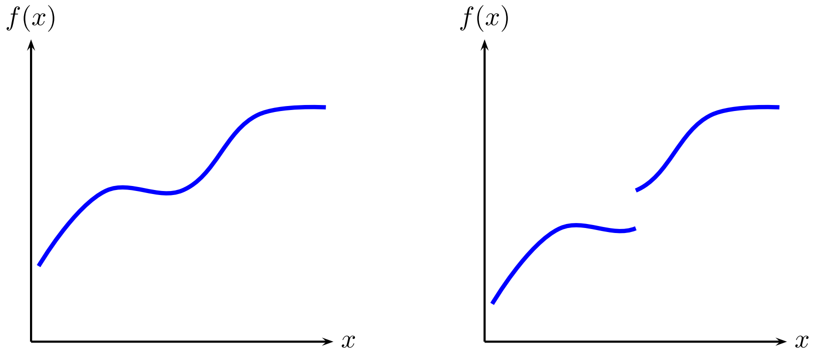


Figure 3.1 Probably not always the best way to envision continuity versus discontinuity

possible for a function continuous at a point to have a graph with lots of "gaps" in a neighbourhood of that point. Thus the "graph gap" characterisation of continuity is a little misleading.

Let us give some examples of functions that are continuous or not. More examples of discontinuous functions are given in Example 3.1.9 below. We suppose the reader to be familiar with the usual collection of "standard functions," at least for the moment. We shall consider some such functions in detail in Section 3.6.

3.1.2 Examples (Continuous and discontinuous functions)

1. For $\alpha \in \mathbb{R}$, define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = \alpha$. Since $|f(x) - f(x_0)| = 0$ for all $x, x_0 \in \mathbb{R}$, it follows immediately that f is continuous.
2. Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = x$. For $x_0 \in \mathbb{R}$ and $\epsilon \in \mathbb{R}_{>0}$ take $\delta = \epsilon$. It then follows that if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \epsilon$, giving continuity of f.
3. Define $f\colon \mathbb{R} \to \mathbb{R}$ by
$$f(x) = \begin{cases} x \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$
We claim that f is continuous. We first note that the functions $f_1, f_2\colon \mathbb{R} \to \mathbb{R}$ defined by
$$f_1(x) = x, \quad f_2(x) = \sin x$$
are continuous. Indeed, f_1 is continuous from part 2 and in Section 3.6 we will prove that f_2 is continuous. The function $f_3\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ defined by $f_3(x) = \frac{1}{x}$ is continuous on any interval not containing 0 by Proposition 3.1.15 below. It then follows from Propositions 3.1.15 and 3.1.16 below that f is continuous at x_0, provided that $x_0 \neq 0$. To show continuity at $x = 0$, let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \epsilon$. Then, provided that $|x| < \delta$,
$$|f(x) - f(0)| = \left|x \sin \tfrac{1}{x}\right| \leq |x| < \epsilon,$$
using the fact that image(sin) $\subseteq [-1, 1]$. This shows that f is continuous at 0, and so is continuous.

4. Define $f\colon \mathbb{R} \to \mathbb{R}$ by
$$f(x) = \begin{cases} x, & x \in \mathbb{Q}, \\ 0, & \text{otherwise}. \end{cases}$$
We claim that f is continuous at $x_0 = 0$ and discontinuous everywhere else.
To see that f is continuous at $x_0 = 0$, let $\epsilon \in \mathbb{R}_{>0}$ and choose $\delta = \epsilon$. Then, for $|x - x_0| < \delta$ we have either $f(x) = x$ or $f(x) = 0$. In either case, $|f(x) - f(x_0)| < \epsilon$, showing that f is indeed continuous at $x_0 = 0$. Note that this is a function whose continuity at $x_0 = 0$ is not subject to an interpretation like that of Figure 3.1 since the graph of f has an uncountable number of "gaps" near 0.
Next we show that f is discontinuous at x_0 for $x_0 \neq 0$. We have two possibilities.
   (a) $x_0 \in \mathbb{Q}$: Let $\epsilon < \frac{1}{2}|x_0|$. For any $\delta \in \mathbb{R}_{>0}$ the set $\mathsf{B}(\delta, x_0)$ will contain points $x \in \mathbb{R}$ for which $f(x) = 0$. Thus for any $\delta \in \mathbb{R}_{>0}$ the set $\mathsf{B}(\delta, x_0)$ will contain points x such that $|f(x) - f(x_0)| = |x_0| > \epsilon$. This shows that f is discontinuous at nonzero rational numbers.
   (b) $x_0 \in \mathbb{R} \setminus \mathbb{Q}$: Let $\epsilon = \frac{1}{2}|x_0|$. For any $\delta \in \mathbb{R}_{>0}$ we claim that the set $\mathsf{B}(\delta, x_0)$ will contain points $x \in \mathbb{R}$ for which $|f(x)| > \epsilon$ (why?). It then follows that for any $\delta \in \mathbb{R}_{>0}$ the set $\mathsf{B}(\delta, x_0)$ will contain points x such that $|f(x) - f(x_0)| = |f(x)| > \epsilon$, so showing that f is discontinuous at all irrational numbers.
5. Let $I = (0, \infty)$ and on I define the function $f\colon I \to \mathbb{R}$ by $f(x) = \frac{1}{x}$. It follows from Proposition 3.1.15 below that f is continuous on I.
6. Next take $I = [0, \infty)$ and define $f\colon I \to \mathbb{R}$ by
$$f(x) = \begin{cases} \frac{1}{x}, & x \in \mathbb{R}_{>0}, \\ 0, & x = 0. \end{cases}$$
In the previous example we saw that f is continuous at all points in $(0, \infty)$. However, at $x = 0$ the function is discontinuous, as is easily verified. •

The following alternative characterisations of continuity are sometimes useful. The first of these, part (ii) in the theorem, will also be helpful in motivating the general definition of continuity given for topological spaces in Section **??**. The reader will wish to recall from Notation 2.3.28 the notation $\lim_{x \to_I x_0} f(x)$ for taking limits in intervals.

3.1.3 Theorem (Alternative characterisations of continuity) *For a function* $f\colon I \to \mathbb{R}$ *defined on an interval* I *and for* $x_0 \in I$*, the following statements are equivalent:*

(i) f *is continuous at* x_0*;*

(ii) for every neighbourhood V *of* $f(x_0)$ *there exists a neighbourhood* U *of* x_0 *in* I *such that* $f(U) \subseteq V$*;*

(iii) $\lim_{x \to_I x_0} f(x) = f(x_0)$.

Proof (i) $\Longrightarrow$ (ii) Let $V \subseteq \mathbb{R}$ be a neighbourhood of $f(x_0)$. Let $\epsilon \in \mathbb{R}_{>0}$ be defined such that $\mathsf{B}(\epsilon, f(x_0)) \subseteq V$, this being possible since V is open. Since f is continuous at x_0, there exists $\delta \in \mathbb{R}_{>0}$ such that, if $x \in \mathsf{B}(\delta, x_0) \cap I$, then we have $f(x) \in \mathsf{B}(\epsilon, f(x_0))$. This shows that, around the point x_0, we can find an open set in I whose image lies in V.

(ii) $\Longrightarrow$ (iii) Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in I converging to x_0 and let $\epsilon \in \mathbb{R}_{>0}$. By hypothesis there exists a neighbourhood U of x_0 in I such that $f(U) \subseteq \mathsf{B}(\epsilon, f(x_0))$. Thus there exists $\delta \in \mathbb{R}_{>0}$ such that $f(\mathsf{B}(\delta, x_0) \cap I) \subseteq \mathsf{B}(\epsilon, f(x_0))$ since U is open in I. Now choose $N \in \mathbb{Z}_{>0}$ sufficiently large that $|x_j - x_0| < \delta$ for $j \geq N$. It then follows that $|f(x_j) - f(x_0)| < \epsilon$ for $j \geq N$, so giving convergence of $(f(x_j))_{j\in\mathbb{Z}_{>0}}$ to $f(x_0)$, as desired, after an application of Proposition 2.3.29.

(iii) $\Longrightarrow$ (i) Let $\epsilon \in \mathbb{R}_{>0}$. Then, by definition of $\lim_{x\to_I x_0} f(x) = f(x_0)$, there exists $\delta \in \mathbb{R}_{>0}$ such that, for $x \in \mathsf{B}(\delta, x_0) \cap I$, $|f(x) - f(x_0)| < \epsilon$, which is exactly the definition of continuity of f at x_0. ■

3.1.4 Corollary *For an interval* $\mathrm{I} \subseteq \mathbb{R}$, *a function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is continuous if and only if* $\mathrm{f}^{-1}(\mathrm{V})$ *is open in* I *for every open subset* V *of* $\mathbb{R}$.

Proof Suppose that f is continuous. If $V \cap \mathrm{image}(f) = \emptyset$ then clearly $f^{-1}(V) = \emptyset$ which is open. So assume that $V \cap \mathrm{image}(f) \neq \emptyset$ and let $x \in f^{-1}(V)$. Since f is continuous at x and since V is a neighbourhood of $f(x)$, there exists a neighbourhood U of x such that $f(U) \subseteq V$. Thus $U \subseteq f^{-1}(V)$, showing that $f^{-1}(V)$ is open.

Now suppose that $f^{-1}(V)$ is open for each open set V and let $x \in \mathbb{R}$. If V is a neighbourhood of $f(x)$ then $f^{-1}(V)$ is open. Then there exists a neighbourhood U of x such that $U \subseteq f^{-1}(V)$. By Proposition 1.3.5 we have $f(U) \subseteq f(f^{-1}(V)) \subseteq V$, thus showing that f is continuous. ■

The reader can explore these alternative representations of continuity in Exercise 3.1.9.

A stronger notion of continuity is sometimes useful. As well, the following definition introduces for the first time the important notion of "uniform."

3.1.5 Definition (Uniform continuity) Let $I \subseteq \mathbb{R}$ be an interval. A map $f\colon I \to \mathbb{R}$ is ***uniformly continuous*** if, for every $\epsilon \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ such that $|f(x_1) - f(x_2)| < \epsilon$ whenever $x_1, x_2 \in I$ satisfy $|x_1 - x_2| < \delta$. •

3.1.6 Remark (On the idea of "uniformly") In the preceding definition we have encountered for the first time the idea of a property holding "uniformly." This is an important idea that comes up often in mathematics. Moreover, it is an idea that is often useful in applications of mathematics, since the absence of a property holding "uniformly" can have undesirable consequences. Therefore, we shall say some things about this here.

In fact, the comparison of continuity versus uniform continuity is a good one for making clear the character of something holding "uniformly." Let us compare the definitions.

1. One defines continuity of a function at a point x_0 by asking that, for each $\epsilon \in \mathbb{R}_{>0}$, one can find $\delta \in \mathbb{R}_{>0}$ such that if x is within δ of x_0, then $f(x)$ is within ϵ of $f(x_0)$. Note that δ will generally depend on ϵ, and most importantly for our discussion here, on x_0. Often authors explicitly write $\delta(\epsilon, x_0)$ to denote this dependence of δ on both ϵ and x_0.

2. One defines uniform continuity of a function on the interval I by asking that, for each $\epsilon \in \mathbb{R}_{>0}$, one can find $\delta \in \mathbb{R}_{>0}$ such that if x_1 and x_2 are within δ of one another, then $f(x_1)$ and $f(x_2)$ are within ϵ of one another. Here, the number δ depends *only* on ϵ. Again, to reflect this, some authors explicitly write $\delta(\epsilon)$, or state explicitly that δ is independent of x.

The idea of "uniform" then is that a property, in this case the existence of $\delta \in \mathbb{R}_{>0}$ with a certain property, holds for the entire set I, and not just for a single point. •

Let us give an example to show that uniformly continuous is not the same as continuous.

3.1.7 Example (Uniform continuity versus continuity) Let us give an example of a function that is continuous, but not uniformly continuous. Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. We first show that f is continuous at each point $x_0 \in \mathbb{R}$. Let $\epsilon \in \mathbb{R}_{>0}$ and choose δ such that $2|x_0|\delta + \delta^2 < \epsilon$ (why is this possible?). Then, provided that $|x - x_0| < \delta$, we have

$$\begin{aligned}|f(x) - f(x_0)| &= |x^2 - x_0^2| = |x - x_0||x + x_0| \\ &\leq |x - x_0|(|x| + |x_0|) \leq |x - x_0|(2|x_0| + |x - x_0|) \\ &\leq \delta(2|x_0| + \delta) < \epsilon.\end{aligned}$$

Thus f is continuous.

Now let us show that f is not uniformly continuous. We will show that there exists $\epsilon \in \mathbb{R}_{>0}$ such that there is no $\delta \in \mathbb{R}_{>0}$ for which $|x - x_0| < \delta$ ensures that $|f(x) - f(x_0)| < \epsilon$ *for all* x_0. Let us take $\epsilon = 1$ and let $\delta \in \mathbb{R}_{>0}$. Then define $x_0 \in \mathbb{R}$ such that $\frac{\delta}{2}\left|2x_0 + \frac{\delta}{2}\right| > 1$ (why is this possible?). We then note that $x = x_0 + \frac{\delta}{2}$ satisfies $|x - x_0| < \delta$, but that

$$|f(x) - f(x_0)| = |x^2 - x_0^2| = |x - x_0||x + x_0| = \tfrac{\delta}{2}\left|2x_0 + \tfrac{\delta}{2}\right| > 1 = \epsilon.$$

This shows that f is not uniformly continuous. •

3.1.2 Discontinuous functions[2]

It is often useful to be specific about the nature of a discontinuity of a function that is not continuous. The following definition gives names to all possibilities. The reader may wish to recall from Section 2.3.7 the discussion concerning taking limits using an index set that is a subset of $\mathbb{R}$.

3.1.8 Definition (Types of discontinuity) Let $I \subseteq \mathbb{R}$ be an interval and suppose that $f\colon I \to \mathbb{R}$ is discontinuous at $x_0 \in I$. The point x_0 is:

(i) a ***removable discontinuity*** if $\lim_{x \to_I x_0} f(x)$ exists;

(ii) a ***discontinuity of the first kind***, or a ***jump discontinuity***, if the limits $\lim_{x \downarrow x_0} f(x)$ and $\lim_{x \uparrow x_0} f(x)$ exist;

[2] This section is rather specialised and technical and so can be omitted until needed. However, the material is needed at certain points in the text.

(iii) a ***discontinuity of the second kind***, or an ***essential discontinuity***, if at least one of the limits $\lim_{x\downarrow x_0} f(x)$ and $\lim_{x\uparrow x_0} f(x)$ does not exist.

The set of all discontinuities of f is denoted by D_f. •

In Figure 3.2 we depict the various sorts of discontinuity. We can also illustrate these with explicit examples.

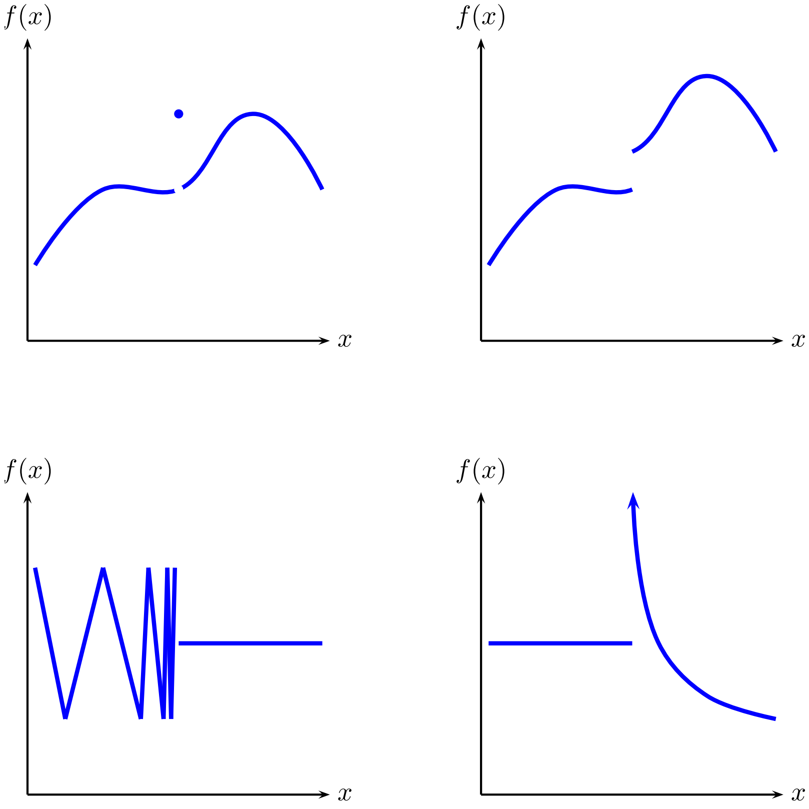


Figure 3.2 A removable discontinuity (top left), a jump discontinuity (top right), and two essential discontinuities (bottom)

3.1.9 Examples (Types of discontinuities)

1. Let $I = [0, 1]$ and let $f \colon I \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x, & x \in (0, 1], \\ 1, & x = 0. \end{cases}$$

It is clear that f is continuous for all $x \in (0, 1]$, and is discontinuous at $x = 0$. However, since we have $\lim_{x\to_I 0} f(x) = 0$ (note that the requirement that this limit be taken in I amounts to the fact that the limit is given by $\lim_{x\downarrow 0} f(x) = 0$), it follows that the discontinuity is removable.

Note that one might be tempted to also say that the discontinuity is a jump discontinuity since the limit $\lim_{x\downarrow 0} f(x)$ exists and since the limit $\lim_{x\uparrow 0} f(x)$

cannot be defined here since 0 is a left endpoint for I. However, we do require that both limits exist at a jump discontinuity, which has as a consequence the fact that jump discontinuities can only occur at interior points of an interval.

2. Let $I = [-1,1]$ and define $f: I \to \mathbb{R}$ by $f(x) = \operatorname{sign}(x)$. We may easily see that f is continuous at $x \in [-1,1] \setminus \{0\}$, and is discontinuous at $x = 0$. Then, since we have $\lim_{x\downarrow 0} f(x) = 1$ and $\lim_{x\uparrow 0} f(x) = -1$, it follows that the discontinuity at 0 is a jump discontinuity.
3. Let $I = [-1,1]$ and define $f: I \to \mathbb{R}$ by
$$f(x) = \begin{cases} \sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$
Then, by Proposition 3.1.15 (and accepting continuity of sin), f is continuous at $x \in [-1,1] \setminus \{0\}$. At $x = 0$ we claim that we have an essential discontinuity. To see this we note that, for any $\epsilon \in \mathbb{R}_{>0}$, the function f restricted to $[0,\epsilon)$ and $(-\epsilon,0]$ takes all possible values in set $[-1,1]$. This is easily seen to preclude existence of the limits $\lim_{x\downarrow 0} f(x)$ and $\lim_{x\uparrow 0} f(x)$.
4. Let $I = [-1,1]$ and define $f: I \to \mathbb{R}$ by
$$f(x) = \begin{cases} \frac{1}{x}, & x \in (0,1], \\ 0, & x \in [-1,0]. \end{cases}$$
Then f is continuous at $x \in [-1,1] \setminus \{0\}$ by Proposition 3.1.15. At $x = 0$ we claim that f has an essential discontinuity. Indeed, we have $\lim_{x\downarrow} f(x) = \infty$, which precludes f having a removable or jump discontinuity at $x = 0$. •

The following definition gives a useful quantitative means of measuring the discontinuity of a function.

3.1.10 Definition (Oscillation) Let $I \subseteq \mathbb{R}$ be an interval and let $f: I \to \mathbb{R}$ be a function. The ***oscillation*** of f is the function $\omega_f: I \to \mathbb{R}$ defined by
$$\omega_f(x) = \inf\{\sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\} \mid \delta \in \mathbb{R}_{>0}\}. \quad \bullet$$

Note that the definition makes sense since the function
$$\delta \mapsto \sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\}$$
is monotonically increasing (see Definition 3.1.27 for a definition of monotonically increasing in this context). In particular, if f is bounded (see Definition 3.1.20 below) then ω_f is also bounded. The following result indicates in what way ω_f measures the continuity of f.

3.1.11 Proposition (Oscillation measures discontinuity) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$ *and a function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$*,* f *is continuous at* $\mathrm{x} \in \mathrm{I}$ *if and only if* $\omega_{\mathrm{f}}(\mathrm{x}) = 0$*.*

Proof Suppose that f is continuous at x and let $\epsilon \in \mathbb{R}_{>0}$. Choose $\delta \in \mathbb{R}_{>0}$ such that if $y \in \mathsf{B}(\delta, x) \cap I$ then $|f(y) - f(x)| < \frac{\epsilon}{2}$. Then, for $x_1, x_2 \in \mathsf{B}(\delta, x)$ we have

$$|f(x_1) - f(x_2)| \le |f(x_1) - f(x)| + |f(x) - f(x_2)| < \epsilon.$$

Therefore,

$$\sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\} < \epsilon.$$

Since ϵ is arbitrary this gives

$$\inf\{\sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\} \mid \delta \in \mathbb{R}_{>0}\} = 0,$$

meaning that $\omega_f(x) = 0$.

Now suppose that $\omega_f(x) = 0$. For $\epsilon \in \mathbb{R}_{>0}$ let $\delta \in \mathbb{R}_{>0}$ be chosen such that

$$\sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\} < \epsilon.$$

In particular, $|f(y) - f(x)| < \epsilon$ for all $y \in \mathsf{B}(\delta, x) \cap I$, giving continuity of f at x. ■

Let us consider a simple example.

3.1.12 Example (Oscillation for a discontinuous function) We let $I = [-1, 1]$ and define $f\colon I \to \mathbb{R}$ by $f(x) = \operatorname{sign}(x)$. It is then easy to see that

$$\omega_f(x) = \begin{cases} 0, & x \neq 0, \\ 2, & x = 0. \end{cases} \qquad \bullet$$

We close this section with a technical property of the oscillation of a function. This property will be useful during the course of some proofs in the text.

3.1.13 Proposition (Closed preimages of the oscillation of a function) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval and let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be a function. Then, for every* $\alpha \in \mathbb{R}_{\geq 0}$*, the set*

$$\mathrm{A}_\alpha = \{\mathrm{x} \in \mathrm{I} \mid \omega_{\mathrm{f}}(\mathrm{x}) \geq \alpha\}$$

is closed in I*.*

Proof The result where $\alpha = 0$ is clear, so we assume that $\alpha \in \mathbb{R}_{>0}$. For $\delta \in \mathbb{R}_{>0}$ define

$$\omega_f(x, \delta) = \sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\}$$

so that $\omega_f(x) = \lim_{\delta \to 0} \omega_f(x, \delta)$. Let $(x_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in A_α converging to $x \in \mathbb{R}$ and let $(\epsilon_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in $(0, \alpha)$ converging to zero. Let $j \in \mathbb{Z}_{>0}$. We claim that there exists points $y_j, z_j \in \mathsf{B}(\epsilon_j, x_j) \cap I$ such that $|f(y_j) - f(z_j)| \geq \alpha - \epsilon_j$. Suppose otherwise so that for every $y, z \in \mathsf{B}(\epsilon_j, x_j) \cap I$ we have $|f(y) - f(z)| < \alpha - \epsilon_j$. It then follows that $\lim_{\delta \to 0} \omega_f(x_j, \delta) \leq \alpha - \epsilon_j < \alpha$, contradicting the fact that $x_j \in A_\alpha$. We claim that $(y_j)_{j \in \mathbb{Z}_{>0}}$ and $(z_j)_{j \in \mathbb{Z}_{>0}}$ converge to x. Indeed, let $\epsilon \in \mathbb{R}_{>0}$ and choose $N_1 \in \mathbb{Z}_{>0}$ sufficiently large that $\epsilon_j < \frac{\epsilon}{2}$ for $j \geq N_1$ and choose $N_2 \in \mathbb{Z}_{>0}$ such that $|x_j - x| < \frac{\epsilon}{2}$ for $j \geq N_2$. Then, for $j \geq \max\{N_1, N_2\}$ we have

$$|y_j - x| \leq |y_j - x_j| + |x_j - x| < \epsilon.$$

Thus $(y_j)_{j\in\mathbb{Z}_{>0}}$ converges to x, and the same argument, and therefore the same conclusion, also applies to $(z_j)_{j\in\mathbb{Z}_{>0}}$.

Thus we have sequences of points $(y_j)_{j\in\mathbb{Z}_{>0}}$ and $(z_j)_{j\in\mathbb{Z}_{>0}}$ in I converging to x and a sequence $(\epsilon_j)_{j\in\mathbb{Z}_{>0}}$ in $(0,\alpha)$ converging to zero for which $|f(y_j)-f(z_j)|\geq\alpha-\epsilon_j$. We claim that this implies that $\omega_f(x)\geq\alpha$. Indeed, suppose that $\omega_f(x)<\alpha$. There exists $N\in\mathbb{Z}_{>0}$ such that $\alpha-\epsilon_j>\alpha-\omega_f(x)$ for every $j\geq N$. Therefore,

$$|f(y_j)-f(z_j)|\geq\alpha-\epsilon_j>\alpha-\omega_f(x)$$

for every $j\geq N$. This contradicts the definition of $\omega_f(x)$ since the sequences $(y_j)_{j\in\mathbb{Z}_{>0}}$ and $(z_j)_{j\in\mathbb{Z}_{>0}}$ converge to x.

Now we claim that the sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ converges to x. Let $\epsilon\in\mathbb{R}_{>0}$ and let $N_1\in\mathbb{Z}_{>0}$ be large enough that $|x-y_j|<\frac{\epsilon}{2}$ for $j\geq N_1$ and let $N_2\in\mathbb{Z}_{>0}$ be large enough that $\epsilon_j<\frac{\epsilon}{2}$ for $j\geq N_2$. Then, for $j\geq\max\{N_1,N_2\}$ we have

$$|x-x_j|\leq|x-y_j|+|y_j-x_j|<\epsilon,$$

as desired.

This shows that every sequence in A_α converges to a point in A_α. It follows from Exercise 2.5.2 that A_α is closed. ■

The following corollary is somewhat remarkable, in that it shows that the set of discontinuities of a function cannot be arbitrary.

3.1.14 Corollary (Discontinuities are the countable union of closed sets) *Let* $\mathrm{I}\subseteq\mathbb{R}$ *be an interval and let* $\mathrm{f}\colon\mathrm{I}\to\mathbb{R}$ *be a function. Then the set*

$$\mathrm{D_f}=\{\mathrm{x}\in\mathrm{I}\mid\ \mathrm{f}\textit{ is not continuous at }\mathrm{x}\}$$

is the countable union of closed sets.

Proof This follows immediately from Proposition 3.1.13 after we note that

$$D_f=\cup_{k\in\mathbb{Z}_{>0}}\{x\in I\mid\ \omega_f(x)\geq\tfrac{1}{k}\}.$$ ■

missing stuff

3.1.3 Continuity and operations on functions

Let us consider how continuity behaves relative to simple operations on functions. To do so, we first note that, given an interval I and two functions $f,g\colon I\to\mathbb{R}$, one can define two functions $f+g, fg\colon I\to\mathbb{R}$ by

$$(f+g)(x)=f(x)+g(x),\qquad(fg)(x)=f(x)g(x),$$

respectively. Moreover, if $g(x)\neq 0$ for all $x\in I$, then we define

$$\Big(\frac{f}{g}\Big)(x)=\frac{f(x)}{g(x)}.$$

Thus one can add and multiply $\mathbb{R}$-valued functions using the operations of addition and multiplication in $\mathbb{R}$.

3.1.15 Proposition (Continuity, and addition and multiplication) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$, *if* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *are continuous at* $\mathrm{x}_0 \in \mathrm{I}$, *then both* $\mathrm{f} + \mathrm{g}$ *and* fg *are continuous at* x_0. *If additionally* $\mathrm{g}(\mathrm{x}) \neq 0$ *for all* $\mathrm{x} \in \mathrm{I}$, *then* $\frac{\mathrm{f}}{\mathrm{g}}$ *is continuous at* x_0.

Proof To show that $f + g$ and fg are continuous at x_0 if f and g are continuous at x_0, let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in I converging to x_0. Then, by Theorem 3.1.3 the sequences $(f(x_j))_{j\in\mathbb{Z}_{>0}}$ and $(g(x_j))_{j\in\mathbb{Z}_{>0}}$ converge to $f(x_0)$ and $g(x_0)$, respectively. Then, by Proposition 2.3.23, the sequences $(f(x_j) + g(x_j))_{j\in\mathbb{Z}_{>0}}$ and $(f(x_j)g(x_j))_{j\in\mathbb{Z}_{>0}}$ converge to $f(x_0) + g(x_0)$ and $f(x_0)g(x_0)$, respectively. Then $\lim_{j\to\infty}(f + g)(x_j) = (f + g)(x_0)$ and $\lim_{j\to\infty}(fg)(x_j) = (fg)(x_0)$, and the result follows by Proposition 2.3.29 and Theorem 3.1.3.

Now suppose that $g(x) \neq 0$ for every $x \in I$. Then there exists $\epsilon \in \mathbb{R}_{>0}$ such that $|g(x_0)| > 2\epsilon$. By Theorem 3.1.3 take $\delta \in \mathbb{R}_{>0}$ such that $g(\mathsf{B}(\delta, x_0)) \subseteq \mathsf{B}(\epsilon, g(x_0))$. Thus g is nonzero on the ball $\mathsf{B}(\delta, x_0)$. Now let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{B}(\delta, x_0)$ converging to x_0. Then, as above, the sequences $(f(x_j))_{j\in\mathbb{Z}_{>0}}$ and $(g(x_j))_{j\in\mathbb{Z}_{>0}}$ converge to $f(x_0)$ and $g(x_0)$, respectively. We can now employ Proposition 2.3.23 to conclude that the sequence $\left(\frac{f(x_j)}{g(x_j)}\right)_{j\in\mathbb{Z}_{>0}}$ converges to $\frac{f(x_0)}{g(x_0)}$, and the last part of the result follows by Proposition 2.3.29 and Theorem 3.1.3. ■

3.1.16 Proposition (Continuity and composition) *Let* $\mathrm{I}, \mathrm{J} \subseteq \mathbb{R}$ *be intervals and let* $\mathrm{f}\colon \mathrm{I} \to \mathrm{J}$ *and* $\mathrm{f}\colon \mathrm{J} \to \mathbb{R}$ *be continuous at* $\mathrm{x}_0 \in \mathrm{I}$ *and* $\mathrm{f}(\mathrm{x}_0) \in \mathrm{J}$, *respectively. Then* $\mathrm{g} \circ \mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is continuous at* x_0.

Proof Let W be a neighbourhood of $g \circ f(x_0)$. Since g is continuous at $f(x_0)$ there exists a neighbourhood V of $f(x_0)$ such that $g(V) \subseteq W$. Since f is continuous at x_0 there exists a neighbourhood U of x_0 such that $f(U) \subseteq V$. Clearly $g \circ f(U) \subseteq W$, and the result follows from Theorem 3.1.3. ■

3.1.17 Proposition (Continuity and restriction) *If* $\mathrm{I}, \mathrm{J} \subseteq \mathbb{R}$ *are intervals for which* $\mathrm{J} \subseteq \mathrm{I}$, *and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is continuous at* $\mathrm{x}_0 \in \mathrm{J} \subseteq \mathrm{I}$, *then* $\mathrm{f}|\mathrm{J}$ *is continuous at* x_0.

Proof This follows immediately from Theorem 3.1.3, also using Proposition 1.3.5, after one notes that open subsets of J are of the form $U \cap I$ where U is an open subset of I. ■

Note that none of the proofs of the preceding results use the definition of continuity, but actually use the alternative characterisations of Theorem 3.1.3. Thus these alternative characterisations, while less intuitive initially (particularly the one involving open sets), they are in fact quite useful.

Let us finally consider the behaviour of continuity with respect to the operations of selection of maximums and minimums.

3.1.18 Proposition (Continuity and min and max) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *are continuous functions, then the functions*

$$\mathrm{I} \ni \mathrm{x} \mapsto \min\{\mathrm{f}(\mathrm{x}), \mathrm{g}(\mathrm{x})\} \in \mathbb{R}, \qquad \mathrm{I} \ni \mathrm{x} \mapsto \max\{\mathrm{f}(\mathrm{x}), \mathrm{g}(\mathrm{x})\} \in \mathbb{R}$$

are continuous.

Proof Let $x_0 \in I$ and let $\epsilon \in \mathbb{R}_{>0}$. Let us first assume that $f(x_0) > g(x_0)$. That is to say, assume that $(f - g)(x_0) \in \mathbb{R}_{>0}$. Continuity of f and g ensures that there exists $\delta_1 \in \mathbb{R}_{>0}$ such that if $x \in \mathsf{B}(\delta_1, x_0) \cap I$ then $(f - g)(x) \in \mathbb{R}_{>0}$. That is, if $x \in \mathsf{B}(\delta_1, x_0) \cap I$ then

$$\min\{f(x), g(x)\} = g(x), \quad \max\{f(x), g(x)\} = f(x).$$

Continuity of f ensures that there exists $\delta_2 \in \mathbb{R}_{>0}$ such that if $x \in \mathsf{B}(\delta_2, x_0) \cap I$ then $|f(x) - f(x_0)| < \epsilon$. Similarly, continuity of f ensures that there exists $\delta_3 \in \mathbb{R}_{>0}$ such that if $x \in \mathsf{B}(\delta_3, x_0) \cap I$ then $|g(x) - g(x_0)| < \epsilon$. Let $\delta_4 = \min\{\delta_1, \delta_2\}$. If $x \in \mathsf{B}(\delta_4, x_0) \cap I$ then

$$|\min\{f(x), g(x)\} - \min\{f(x_0), g(x_0)\}| = |g(x) - g(x_0)| < \epsilon$$

and

$$|\max\{f(x), g(x)\} - \max\{f(x_0), g(x_0)\}| = |f(x) - f(x_0)| < \epsilon.$$

This gives continuity of the two functions in this case. Similarly, swapping the rôle of f and g, if $f(x_0) < g(x_0)$ one can arrive at the same conclusion. Thus we need only consider the case when $f(x_0) = g(x_0)$. In this case, by continuity of f and g, choose $\delta \in \mathbb{R}_{>0}$ such that $|f(x) - f(x_0)| < \epsilon$ and $|g(x) - g(x_0)| < \epsilon$ for $x \in \mathsf{B}(\delta, x_0) \cap I$. Then let $x \in \mathsf{B}(\delta, x_0) \cap I$. If $f(x) \geq g(x)$ then we have

$$|\min\{f(x), g(x)\} - \min\{f(x_0), g(x_0)\}| = |g(x) - g(x_0)| < \epsilon$$

and

$$|\max\{f(x), g(x)\} - \max\{f(x_0), g(x_0)\}| = |f(x) - f(x_0)| < \epsilon.$$

This gives the result in this case, and one similarly gets the result when $f(x) < g(x)$. ■

3.1.4 Continuity, and compactness and connectedness

In this section we will consider some of the relationships that exist between continuity, and compactness and connectedness. We see here for the first time some of the benefits that can be drawn from the notion of continuity. Moreover, if one studies the proofs of the results in this section, one can see that we use the actual definition of compactness (rather than the simpler alternative characterisation of compact sets as being closed and bounded) to great advantage.

The first result is a simple and occasionally useful one.

3.1.19 Proposition (The continuous image of a compact set is compact) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is a compact interval and if* $\mathrm{f}\colon \mathrm{I} \rightarrow \mathbb{R}$ *is continuous, then* $\mathrm{image(f)}$ *is compact.*

Proof Let $(U_a)_{a \in A}$ be an open cover of $\mathrm{image}(f)$. Then $(f^{-1}(U_a))_{a \in A}$ is an open cover of I, and so there exists a finite subset $(a_1, \dots, a_k) \subseteq A$ such that $\cup_{j=1}^{k} f^{-1}(U_{a_k}) = I$. It is then clear that $(f(f^{-1}(U_{a_1})), \dots, f(f^{-1}(U_{a_k})))$ covers $\mathrm{image}(f)$. Moreover, by Proposition 1.3.5, $f(f^{-1}(U_{a_j})) \subseteq U_{a_j}$, $j \in \{1, \dots, k\}$. Thus $(U_{a_1}, \dots, U_{a_k})$ is a finite subcover of $(U_a)_{a \in A}$. ■

A useful feature that a function might possess is that of having bounded values.

3.1.20 Definition (Bounded function) For an interval I, a function $f\colon I \rightarrow \mathbb{R}$ is:

(i) ***bounded*** if there exists $M \in \mathbb{R}_{>0}$ such that $\mathrm{image}(f) \subseteq \overline{\mathsf{B}}(M, 0)$;

(ii) ***locally bounded*** if $f|J$ is bounded for every compact interval $J \subseteq I$;

(iii) ***unbounded*** if it is not bounded. •

3.1.21 Remark (On "locally") This is our first encounter with the qualifier "locally" assigned to a property, in this case, of a function. This concept will appear frequently, as for example in this chapter with the notion of "locally bounded variation" (Definition 3.3.6) and "locally absolutely continuous" (Definition 5.9.23). The idea in all cases is the same; that a property holds "locally" if it holds on every compact subset. •

For continuous functions it is sometimes possible to immediately assert boundedness simply from the property of the domain.

3.1.22 Theorem (Continuous functions on compact intervals are bounded) *If* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *is a compact interval, then a continuous function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is bounded.*

Proof Let $x \in I$. As f is continuous, there exists $\delta \in \mathbb{R}_{>0}$ so that $|f(y) - f(x)| < 1$ provided that $|y - x| < \delta$. In particular, if $x \in I$, there is an open interval I_x in I with $x \in I_x$ such that $|f(y)| \leq |f(x)| + 1$ for all $x \in I_x$. Thus f is bounded on I_x. This can be done for each $x \in I$, so defining a family of open sets $(I_x)_{x\in I}$. Clearly $I \subseteq \cup_{x\in I} I_x$, and so, by Theorem 2.5.27, there exists a finite collection of points $x_1, \dots, x_k \in I$ such that $I \subseteq \cup_{j=1}^{k} I_{x_j}$. Obviously for any $x \in I$,

$$|f(x)| \leq 1 + \max\{f(x_1), \dots, f(x_k)\},$$

thus showing that f is bounded. ■

In Exercise 3.1.7 the reader can explore cases where the theorem does not hold. Related to the preceding result is the following.

3.1.23 Theorem (Continuous functions on compact intervals achieve their extreme values) *If* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *is a compact interval and if* $\mathrm{f}\colon [\mathrm{a}, \mathrm{b}] \to \mathbb{R}$ *is continuous, then there exist points* $\mathrm{x}_{\mathrm{min}}, \mathrm{x}_{\mathrm{max}} \in [\mathrm{a}, \mathrm{b}]$ *such that*

$$\mathrm{f}(\mathrm{x}_{\mathrm{min}}) = \inf\{\mathrm{f}(\mathrm{x}) \mid \mathrm{x} \in [\mathrm{a}, \mathrm{b}]\}, \quad \mathrm{f}(\mathrm{x}_{\mathrm{max}}) = \sup\{\mathrm{f}(\mathrm{x}) \mid \mathrm{x} \in [\mathrm{a}, \mathrm{b}]\}.$$

Proof It suffices to show that f achieves its maximum on I since if f achieves its maximum, then $-f$ will achieve its minimum. So let $M = \sup\{f(x) \mid x \in I\}$, and suppose that there is no point $x_{\max} \in I$ for which $f(x_{\max}) = M$. Then $f(x) < M$ for each $x \in I$. For a given $x \in I$ we have

$$f(x) = \tfrac{1}{2}(f(x) + f(x)) < \tfrac{1}{2}(f(x) + M).$$

Continuity of f ensures that there is an open interval I_x containing x such that, for each $y \in I_x \cap I$, $f(y) < \frac{1}{2}(f(x) + M)$. Since $I \subseteq \cup_{x\in I} I_x$, by the Heine–Borel theorem, there exists a finite number of points $x_1, \dots, x_k$ such that $I \subseteq \cup_{j=1}^{k} I_{x_j}$. Let $m = \max\{f(x_1), \dots, f(x_k)\}$ so that, for each $y \in I_{x_j}$, and for each $j \in \{1, \dots, k\}$, we have

$$f(y) < \tfrac{1}{2}(f(x_j) + M) < \tfrac{1}{2}(m + M),$$

which shows that $\frac{1}{2}(m + M)$ is an upper bound for f. However, since f attains the value m on I, we have $m < M$ and so $\frac{1}{2}(m + M) < M$, contradicting the fact that M is the least upper bound. Thus our assumption that f cannot attain the value M on I is false. ■

The theorem tells us that a continuous function on a bounded interval actually *attains* its maximum and minimum value *on the interval*. You should understand that this is not the case if I is neither closed nor bounded (see Exercise 3.1.8).

Our next result gives our first connection between the concepts of uniformity and compactness. This is something of a theme in analysis where continuity is involved. A good place to begin to understand the relationship between compactness and uniformity is the proof of the following theorem, since it is one of the simplest instances of the phenomenon.

3.1.24 Theorem (Heine–Cantor Theorem) *Let* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *be a compact interval. If* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is continuous, then it is uniformly continuous.*

Proof Let $x \in [a, b]$ and let $\epsilon \in \mathbb{R}_{>0}$. Since f is continuous, then there exists $\delta_x \in \mathbb{R}_{>0}$ such that, if $|y - x| < \delta_x$, then $|f(y) - f(x)| < \frac{\epsilon}{2}$. Now define an open interval $I_x = (x - \frac{1}{2}\delta_x, x + \frac{1}{2}\delta_x)$. Note that $[a, b] \subseteq \cup_{x\in[a,b]} I_x$, so that the open sets $(I_x)_{x\in[a,b]}$ cover $[a, b]$. By definition of compactness, there then exists a finite number of open sets from $(I_x)_{x\in[a,b]}$ that cover $[a, b]$. Denote this finite family by $(I_{x_1}, \dots, I_{x_k})$ for some $x_1, \dots, x_k \in [a, b]$. Take $\delta = \frac{1}{2}\min\{\delta_{x_1}, \dots, \delta_{x_k}\}$. Now let $x, y \in [a, b]$ satisfy $|x - y| < \delta$. Then there exists $j \in \{1, \dots, k\}$ such that $x \in I_{x_j}$ since the sets $I_{x_1}, \dots, I_{x_k}$ cover $[a, b]$. We also have

$$|y - x_j| = |y - x + x - x_j| \le |y - x| + |x - x_j| < \tfrac{1}{2}\delta_{x_j} + \tfrac{1}{2}\delta_{x_j} = \delta_{x_j},$$

using the triangle inequality. Therefore,

$$\begin{aligned}|f(y) - f(x)| &= |f(y) - f(x_j) + f(x_j) - f(x)| \\ &\le |f(y) - f(x_j)| + |f(x_j) - f(x)| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,\end{aligned}$$

again using the triangle inequality. Since this holds for *any* $x \in [a, b]$, it follows that f is uniformly continuous. ■

Next we give a standard result from calculus that is frequently useful.

3.1.25 Theorem (Intermediate Value Theorem) *Let* I *be an interval and let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be continuous. If* $\mathrm{x}_1, \mathrm{x}_2 \in \mathrm{I}$ *then, for any* $\mathrm{y} \in [\mathrm{f}(\mathrm{x}_1), \mathrm{f}(\mathrm{x}_2)]$, *there exists* $\mathrm{x} \in \mathrm{I}$ *such that* $\mathrm{f}(\mathrm{x}) = \mathrm{y}$.

Proof Since otherwise the result is obviously true, we may suppose that $y \in (f(x_1), f(x_2))$. Also, since we may otherwise replace f with $-f$, we may without loss of generality suppose that $x_1 < x_2$. Now define $S = \{x \in [x_1, x_2] \mid f(x) \le y\}$ and let $x_0 = \sup S$. We claim that $f(x_0) = y$. Suppose not. Then first consider the case where $f(x_0) > y$, and define $\epsilon = f(x_0) - y$. Then there exists $\delta \in \mathbb{R}_{>0}$ such that $|f(x) - f(x_0)| < \epsilon$ for $|x - x_0| < \delta$. In particular, $f(x_0 - \delta) > y$, contradicting the fact that $x_0 = \sup S$. Next suppose that $f(x_0) < y$. Let $\epsilon = y - f(x_0)$ so that there exists $\delta \in \mathbb{R}_{>0}$ such that $|f(x) - f(x_0)| < \epsilon$ for $|x - x_0| < \delta$. In particular, $f(x_0 + \delta) < y$, contradicting again the fact that $x_0 = \sup S$. ■

In Figure 3.3 we give the idea of the proof of the Intermediate Value Theorem. There is also a useful relationship between continuity and connected sets.

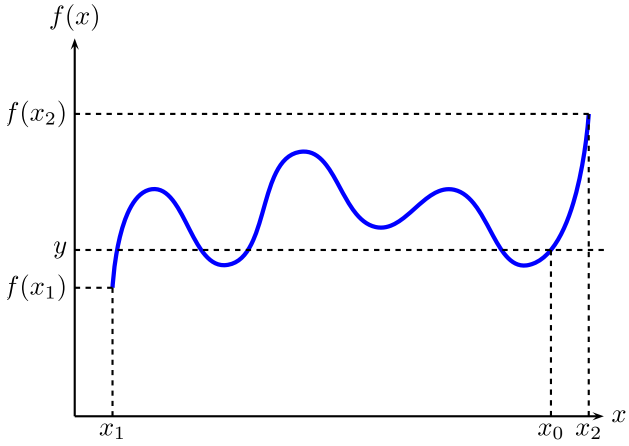


Figure 3.3 Illustration of the Intermediate Value Theorem

3.1.26 Proposition (The continuous image of a connected set is connected) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval, if* $\mathrm{S} \subseteq \mathrm{I}$ *is connected, and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is continuous, then* $\mathrm{f(S)}$ *is connected.*

Proof Suppose that $f(S)$ is not connected. Then there exist nonempty separated sets A and B such that $f(S) = A \cup B$. Let $C = S \cap f^{-1}(A)$ and $D = S \cap f^{-1}(B)$. By Propositions 1.1.4 and 1.3.5 we have

$$\begin{aligned} C \cup D &= (S \cap f^{-1}(A)) \cup (S \cap f^{-1}(B)) \\ &= S \cap (f^{-1}(A) \cup f^{-1}(B)) = S \cap f^{-1}(A \cup B) = S. \end{aligned}$$

By Propositions 2.5.20 and 1.3.5, and since $f^{-1}(\mathrm{cl}(A))$ is closed, we have

$$\mathrm{cl}(C) = \mathrm{cl}(f^{-1}(A)) \subseteq \mathrm{cl}(f^{-1}(\mathrm{cl}(A)) = f^{-1}(\mathrm{cl}(A)).$$

We also clearly have $D \subseteq f^{-1}(B)$. Therefore, by Proposition 1.3.5,

$$\mathrm{cl}(C) \cap D \subseteq f^{-1}(\mathrm{cl}(A)) \cap f^{-1}(B) = f^{-1}(\mathrm{cl}(A) \cap B) = \emptyset.$$

We also similarly have $C \cap \mathrm{cl}(D) = \emptyset$. Thus S is not connected, which gives the result. ■

3.1.5 Monotonic functions and continuity

In this section we consider a special class of functions, namely those that are "increasing" or "decreasing."

3.1.27 Definition (Monotonic function) For $I \subseteq \mathbb{R}$ an interval, a function $f\colon I \to \mathbb{R}$ is:

(i) ***monotonically increasing*** if, for every $x_1, x_2 \in I$ with $x_1 < x_2$, $f(x_1) \leq f(x_2)$;

(ii) ***strictly monotonically increasing*** if, for every $x_1, x_2 \in I$ with $x_1 < x_2$, $f(x_1) < f(x_2)$;

(iii) ***monotonically decreasing*** if, for every $x_1, x_2 \in I$ with $x_1 < x_2$, $f(x_1) \geq f(x_2)$;

(iv) ***strictly monotonically decreasing*** if, for every $x_1, x_2 \in I$ with $x_1 < x_2$, $f(x_1) > f(x_2)$;

(v) ***constant*** if there exists $\alpha \in \mathbb{R}$ such that $f(x) = \alpha$ for every $x \in I$. •

Let us see how monotonicity can be used to make some implications about the continuity of a function. In Theorem 3.2.26 below we will explore some further properties of monotonic functions.

3.1.28 Theorem (Characterisation of monotonic functions I) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is either monotonically increasing or monotonically decreasing, then the following statements hold:*

(i) the limits $\lim_{\mathrm{x}\downarrow\mathrm{x}_0} \mathrm{f}(\mathrm{x})$ *and* $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \mathrm{f}(\mathrm{x})$ *exist whenever they make sense as limits in* I*;*

(ii) the set on which f *is discontinuous is countable.*

Proof We can assume without loss of generality (why?), we assume that $I = [a, b]$ and that f is monotonically increasing.

(i) First let us consider limits from the left. Thus let $x_0 > a$ and consider $\lim_{x\uparrow x_0} f(x)$. For any increasing sequence $(x_j)_{j\in\mathbb{Z}_{>0}} \subseteq [a, x_0)$ converging to x_0 the sequence $(f(x_j))_{j\in\mathbb{Z}_{>0}}$ is bounded and increasing. Therefore it has a limit by Theorem 2.3.8. In a like manner, one shows that right limits also exist.

(ii) Define

$$j(x_0) = \lim_{x\downarrow x_0} f(x) - \lim_{x\uparrow x_0} f(x)$$

as the jump at x_0. This is nonzero if and only if x_0 is a point of discontinuity of f. Let A_f be the set of points of discontinuity of f. Since f is monotonically increasing and defined on a compact interval, it is bounded and we have

$$\sum_{x\in A_f} j(x) \leq f(b) - f(a). \tag{3.1}$$

Now let $n \in \mathbb{Z}_{>0}$ and denote

$$A_n = \left\{x \in [a, b] \;\middle|\; j(x) > \tfrac{1}{n}\right\}.$$

The set A_n must be finite by (3.1). We also have

$$A_f = \bigcup_{n\in\mathbb{Z}_{>0}} A_n,$$

meaning that A_f is a countable union of finite sets. Thus A_f is itself countable. ■

Sometimes the following "local" characterisation of monotonicity is useful.

3.1.29 Proposition (Monotonicity is "local") *A function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *defined on an interval* I *is*

(i) monotonically increasing if and only if, for every $\mathrm{x} \in \mathrm{I}$*, there exists a neighbourhood* U *of* x *such that* $\mathrm{f}|\mathrm{U} \cap \mathrm{I}$ *is monotonically increasing;*

(ii) strictly monotonically increasing if and only if, for every $\mathrm{x} \in \mathrm{I}$*, there exists a neighbourhood* U *of* x *such that* $\mathrm{f}|\mathrm{U} \cap \mathrm{I}$ *is strictly monotonically increasing;*

(iii) monotonically decreasing if and only if, for every $\mathrm{x} \in \mathrm{I}$*, there exists a neighbourhood* U *of* x *such that* $\mathrm{f}|\mathrm{U} \cap \mathrm{I}$ *is monotonically decreasing;*

(iv) strictly monotonically decreasing if and only if, for every $\mathrm{x} \in \mathrm{I}$*, there exists a neighbourhood* U *of* x *such that* $\mathrm{f}|\mathrm{U} \cap \mathrm{I}$ *is strictly monotonically decreasing.*

Proof We shall only prove the first assertion as the other follow from an identical sort of argument. Also, the "only if" assertion is clear, so we need only prove the "if" assertion.

Let $x_1, x_2 \in I$ with $x_1 < x_2$. By hypothesis, for $x \in [x_1, x_2]$, there exists $\epsilon_x \in \mathbb{R}_{>0}$ such that, if we define $U_x = (x - \epsilon, x + \epsilon)$, then $f|U_x \cap I$ is monotonically increasing. Note that $(U_x)_{x \in [x_1, x_2]}$ covers $[x_1, x_2]$ and so, by the Heine–Borel Theorem, there exists $\xi_1, \ldots, \xi_k \in [x_1, x_2]$ such that $[x_1, x_2] \subseteq \cup_{j=1}^k U_{\xi_j}$. We can assume that $\xi_1, \ldots, \xi_k$ are ordered so that $x_1 \in U_{\xi_1}$, that $U_{\xi_{j+1}} \cap U_{\xi_j} \neq \emptyset$, and such that $x_2 \in U_{\xi_k}$. We have that $f|U_{\xi_1} \cap I$ is monotonically increasing. Since $f|U_{\xi_2} \cap I$ is monotonically increasing and since $U_{\xi_1} \cap U_{\xi_2} \neq \emptyset$, we deduce that $f|(U_{\xi_1} \cup U_{\xi_2}) \cap I$ is monotonically increasing. We can continue this process to show that

$$f|(U_{\xi_1} \cup \cdots \cup U_{\xi_k}) \cap I$$

is monotonically increasing, which is the result. ■

In thinking about the graph of a continuous monotonically increasing function, it will not be surprising that there might be a relationship between monotonicity and invertibility. In the next result we explore the precise nature of this relationship.

3.1.30 Theorem (Strict monotonicity and continuity implies invertibility) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval, let* $\mathrm{f} \colon \mathrm{I} \to \mathbb{R}$ *be continuous and strictly monotonically increasing (resp. strictly monotonically decreasing). If* $\mathrm{J} = \mathrm{image(f)}$ *then the following statements hold:*

(i) J *is an interval;*

(ii) there exists a continuous, strictly monotonically increasing (resp. strictly monotonically decreasing) inverse $\mathrm{g} \colon \mathrm{J} \to \mathrm{I}$ *for* f.

Proof We suppose f to be strictly monotonically increasing; the case where it is strictly monotonically decreasing is handled similarly (or follows by considering $-f$, which is strictly monotonically increasing if f is strictly monotonically decreasing).

(i) This follows from Theorem 2.5.34 and Proposition 3.1.26, where it is shown that intervals are the only connected sets, and that continuous images of connected sets are connected.

(ii) Since f is strictly monotonically increasing, if $f(x_1) = f(x_2)$, then $x_1 = x_2$. Thus f is injective as a map from I to J. Clearly $f \colon I \to J$ is also surjective, and so is invertible. Let $y_1, y_2 \in J$ and suppose that $y_1 < y_2$. Then $f(g(y_1)) < f(g(y_2))$, implying that $g(y_1) < g(y_2)$. Thus g is strictly monotonically increasing. It remains to show that the inverse g is continuous. Let $y_0 \in J$ and let $\epsilon \in \mathbb{R}_{>0}$. First suppose that $y_0 \in \mathrm{int}(J)$. Let $x_0 = g(y_0)$ and, supposing ϵ sufficiently small, define $y_1 = f(x_0 - \epsilon)$ and $y_2 = f(x_0 + \epsilon)$. Then let $\delta = \min\{y_0 - y_1, y_2 - y_0\}$. If $y \in \mathsf{B}(\delta, y_0)$ then $y \in (y_1, y_2)$, and since g is strictly monotonically increasing

$$x_0 - \epsilon = g(y_1) < g(y) < g(y_2) = x_0 + \epsilon.$$

Thus $g(y) \in \mathsf{B}(\epsilon, y_0)$, giving continuity of g at x_0. An entirely similar argument can be given if y_0 is an endpoint of J. ■

3.1.6 Convex functions and continuity

In this section we see for the first time the important notion of convexity, here in a fairly simple setting.

Let us first define what we mean by a convex function.

3.1.31 Definition (Convex function) For an interval $I \subseteq \mathbb{R}$, a function $f\colon I \to \mathbb{R}$ is:

(i) ***convex*** if

$$f((1-s)x_1 + sx_2) \leq (1-s)f(x_1) + sf(x_2)$$

for every $x_1, x_2 \in I$ and $s \in [0,1]$;

(ii) ***strictly convex*** if

$$f((1-s)x_1 + sx_2) < (1-s)f(x_1) + sf(x_2)$$

for every distinct $x_1, x_2 \in I$ and for every $s \in (0,1)$;

(iii) ***concave*** if $-f$ is convex;

(iv) ***strictly concave*** if $-f$ is strictly convex. •

Let us give some examples of convex functions.

3.1.32 Examples (Convex functions)

1. A constant function $x \mapsto c$, defined on any interval, is both convex and concave in a trivial way. It is neither strictly convex nor strictly concave.
2. A linear function $x \mapsto ax+b$, defined on any interval, is both convex and concave. It is neither strictly convex nor strictly concave.
3. The function $x \mapsto x^2$, defined on any interval, is strictly convex. Let us verify this. For $s \in (0,1)$ and for $x, y \in \mathbb{R}$ we have, using the triangle inequality,
$$((1-s)x + sy)^2 \leq |(1-s)x + sy|^2 < (1-s)^2x^2 + s^2y^2 \leq (1-s)x^2 + sy^2.$$
4. We refer to Section 3.6.1 for the definition of exponential function $\exp\colon \mathbb{R} \to \mathbb{R}$. We claim that exp is strictly convex. This can be verified explicitly with some effort. However, it follows easily from the fact, proved as Proposition 3.2.30 below, that a function like exp that is twice continuously differentiable with a positive second-derivative is strictly convex. (Note that $\exp'' = \exp$.)
5. We claim that the function log defined in Section 3.6.2 is strictly concave as a function on $\mathbb{R}_{>0}$. Here we compute $\log''(x) = -\frac{1}{x^2}$, which gives strict convexity of $-\log$ (and hence strict concavity of log) by Proposition 3.2.30 below.
6. For $x_0 \in \mathbb{R}$, the function $n_{x_0}\colon \mathbb{R} \to \mathbb{R}$ defined by $n_{x_0} = |x - x_0|$ is convex. Indeed, if $x_1, x_2 \in \mathbb{R}$ and $s \in [0,1]$ then
$$\begin{aligned} n_{x_0}((1-s)x_1 + sx_2) &= |(1-s)x_1 + sx_2 - x_0| = |(1-s)(x_1 - x_0) + s(x_2 - x_0)| \\ &\leq (1-s)|x_1 - x_0| + s|x_2 - x_0| = (1-s)n_{x_0}(x_1) + sn_{x_0}(x_2), \end{aligned}$$
using the triangle inequality. •

Let us give an alternative and insightful characterisation of convex functions. For an interval $I \subseteq \mathbb{R}$ define

$$E_I = \{(x, y) \in I^2 \mid s < t\}$$

and, for $a, b \in I$, denote

$$L_b = \{a \in I \mid (a, b) \in E_I\}, \quad R_a = \{b \in I \mid (a, b) \in E_I\}.$$

Now, for $f\colon I \to \mathbb{R}$ define $s_f\colon E_I \to \mathbb{R}$ by

$$s_f(a, b) = \frac{f(b) - f(a)}{b - a}.$$

With this notation at hand, we have the following result.

3.1.33 Lemma (Alternative characterisation of convexity) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$*, a function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is (strictly) convex if and only if, for every* $\mathrm{a}, \mathrm{b} \in \mathrm{I}$*, the functions*

$$\mathrm{L_b} \ni \mathrm{a} \mapsto \mathrm{s_f}(\mathrm{a}, \mathrm{b}) \in \mathbb{R}, \quad \mathrm{R_a} \ni \mathrm{b} \mapsto \mathrm{s_f}(\mathrm{a}, \mathrm{b}) \in \mathbb{R} \tag{3.2}$$

are (strictly) monotonically increasing.

Proof First suppose that f is convex. Let $a, b, c \in I$ satisfy $a < b < c$. Define $s \in (0, 1)$ by $s = \frac{b-a}{c-a}$ and note that the definition of convexity using this value of s gives

$$f(b) \le \frac{c - b}{c - a} f(a) + \frac{b - a}{c - a} f(c).$$

Simple rearrangement gives

$$\frac{c - b}{c - a} f(a) + \frac{b - a}{c - a} f(c) = f(a) + \frac{f(c) - f(a)}{c - a}(b - a) = f(c) - \frac{f(c) - f(a)}{c - a}(c - b),$$

and so we have

$$\frac{f(b) - f(a)}{b - a} \le \frac{f(c) - f(a)}{c - a}, \quad \frac{f(c) - f(a)}{c - a} \le \frac{f(c) - f(b)}{c - b}.$$

In other words, $s_f(a, b) \le s_f(a, c)$ and $s_f(a, c) \le s_f(b, c)$. Since this holds for every $a, b, c \in I$ with $a < b < c$, we conclude that the functions (3.2) are monotonically increasing, as stated. If f is strictly convex, then the inequalities in the above computation are strict, and one concludes that the functions (3.2) are strictly monotonically increasing.

Next suppose that the functions (3.2) are monotonically increasing and let $a, c \in I$ with $a < c$ and let $s \in (0, 1)$. Define $b = (1 - s)a + sc$. A rearrangement of the inequality $s_f(a, b) \le s_f(a, c)$ gives

$$\begin{aligned} & f(b) \le \frac{c - b}{c - a} f(a) + \frac{b - a}{c - a} f(c) \\ \Longrightarrow \quad & f((1 - s)a + sc) \le (1 - s) f(a) + s f(c), \end{aligned}$$

showing that f is convex since $a, c \in I$ with $a < c$ and $s \in (0, 1)$ are arbitrary in the above computation. If the functions (3.2) are strictly monotonically increasing, then the inequalities in the preceding computations are strict, and so one deduces that f is strictly convex. ■

In Figure 3.4 we depict what the lemma is telling us about convex functions. The idea is that the slope of the line connecting the points $(a, f(a))$ and $(b, f(b))$ in the plane is nondecreasing in a and b.

The following inequality for convex functions is very often useful.

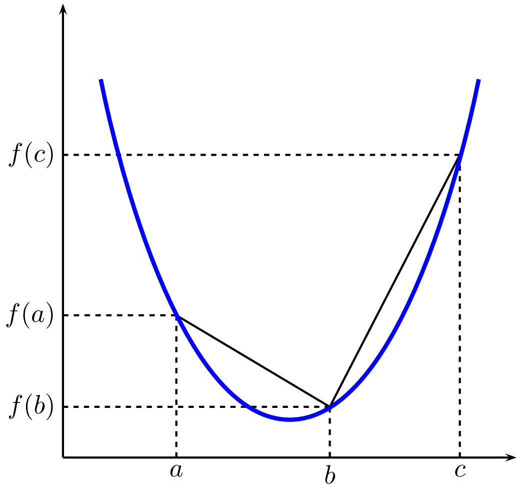


Figure 3.4 A characterisation of a convex function

3.1.34 Theorem (Jensen's inequality) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$*, for a convex function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$*, for* $\mathrm{x}_1, \dots, \mathrm{x}_\mathrm{k} \in \mathrm{I}$*, and for* $\lambda_1, \dots, \lambda_\mathrm{k} \in \mathbb{R}_{\geq 0}$*, we have*

$$\mathrm{f}\Big(\frac{\lambda_1}{\sum_{\mathrm{j}=1}\lambda_\mathrm{j}}\mathrm{x}_1 + \dots + \frac{\lambda_\mathrm{k}}{\sum_{\mathrm{j}=1}^{\mathrm{k}}\lambda_\mathrm{j}}\mathrm{x}_\mathrm{k}\Big) \leq \frac{\lambda_1}{\sum_{\mathrm{j}=1}^{\mathrm{k}}\lambda_\mathrm{j}}\mathrm{f}(\mathrm{x}_1) + \dots + \frac{\lambda_\mathrm{k}}{\sum_{\mathrm{j}=1}^{\mathrm{k}}\lambda_\mathrm{j}}\mathrm{f}(\mathrm{x}_\mathrm{k}).$$

Moreover, if f *is strictly convex and if* $\lambda_1, \dots, \lambda_\mathrm{k} \in \mathbb{R}_{>0}$*, than we have equality in the preceding expression if and only if* $\mathrm{x}_1 = \dots = \mathrm{x}_\mathrm{k}$*.*

Proof We first comment that, with $\lambda_1, \dots, \lambda_k$ and $x_1, \dots, x_k$ as stated,

$$\frac{\lambda_1}{\sum_{j=1}^{k}\lambda_j}x_1 + \dots + \frac{\lambda_k}{\sum_{j=1}^{k}\lambda_j}x_k \in I.$$

This is because intervals are convex, something that will become clear in Section **??**.

It is clear that we can without loss of generality, by replacing λ_j with

$$\lambda'_m = \frac{\lambda_m}{\sum_{j=1}^{k}\lambda_j}, \qquad m \in \{1, \dots, k\},$$

if necessary, that we can assume that $\sum_{j=1}^{k}\lambda_j = 1$.

We first note that if $x_1 = \dots = x_k$ then the inequality in the statement of the theorem is an equality, no matter what the character of f.

The proof is by induction on k, the result being obvious when $k = 1$. So suppose the result is true when $k = m$ and let $x_1, \dots, x_{m+1} \in I$ and let $\lambda_1, \dots, \lambda_{m+1} \in \mathbb{R}_{\geq 0}$ satisfy $\sum_{j=1}^{m+1}\lambda_j = 1$. Without loss of generality (by reindexing if necessary), suppose that $\lambda_{m+1} \in [0, 1)$. Note that

$$\frac{\lambda_1}{1-\lambda_{m+1}} + \dots + \frac{\lambda_m}{1-\lambda_{m+1}} = 1$$

so that, by the induction hypothesis,

$$f\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1 + \dots + \frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big) \leq \frac{\lambda_1}{1-\lambda_{m+1}}f(x_1) + \dots + \frac{\lambda_m}{1-\lambda_{m+1}}f(x_m).$$

Now, by convexity of f,

$$\begin{aligned} f\Big((1-\lambda_{m+1})\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)+\lambda_{m+1}x_{m+1}\Big)& \\ \leq (1-\lambda_{m+1})f\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)&+\lambda_{m+1}f(x_{m+1}). \end{aligned}$$

The desired inequality follows by combining the previous two equations.

To prove the final assertion of the theorem, suppose that f is strictly convex, that $\lambda_1,\dots,\lambda_k\in\mathbb{R}_{>0}$ satisfy $\sum_{j=1}^k\lambda_j=1$, and that the inequality in the theorem is equality. We prove by induction that $x_1=\cdots=x_k$. For $k=1$ the assertion is obvious. Let us prove the assertion for $k=2$. Thus suppose that

$$f((1-\lambda)x_1+\lambda x_2)=(1-\lambda)f(x_1)+\lambda f(x_2)$$

for $x_1,x_2\in I$ and for $\lambda\in(0,1)$. If $x_1\neq x_2$ then we have, by definition of strict convexity,

$$f((1-\lambda)x_1+\lambda x_2)<(1-\lambda)f(x_1)+\lambda f(x_2),$$

contradicting our hypotheses. Thus we must have $x_1=x_2$. Now suppose the assertion is true for $k=m$ and let $x_1,\dots,x_{m+1}\in I$, let $\lambda_1,\dots,\lambda_{m+1}\in\mathbb{R}_{>0}$ satisfy $\sum_{j=1}^{m+1}\lambda_j=1$, and suppose that

$$f(\lambda_1x_1+\cdots+\lambda_{m+1}x_{m+1})=\lambda_1f(x_1)+\cdots+\lambda_{m+1}f(x_{m+1}).$$

Since none of $\lambda_1,\dots,\lambda_{m+1}$ are zero we must have $\lambda_{m+1}\in(0,1)$. Now note that

$$f(\lambda_1x_1+\cdots+\lambda_{m+1}x_{m+1})=f\Big((1-\lambda_{m+1})\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)+\lambda_{m+1}x_{m+1}\Big) \quad (3.3)$$

and that

$$\begin{aligned} \lambda_1f(x_1)+\cdots+\lambda_{m+1}f(x_{m+1})& \\ =(1-\lambda_{m+1})f\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)&+\lambda_{m+1}f(x_{m+1}). \end{aligned}$$

Therefore, by assumption,

$$\begin{aligned} f\Big((1-\lambda_{m+1})\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)+\lambda_{m+1}x_{m+1}\Big)& \\ =(1-\lambda_{m+1})f\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)&+\lambda_{m+1}f(x_{m+1}). \quad (3.4) \end{aligned}$$

Since the assertion we are proving holds for $k=2$ this implies that

$$x_{m+1}=\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m. \quad (3.5)$$

Now suppose that the numbers $x_1,\dots,x_m$ are not all equal. Then, by the induction hypothesis,

$$f\Big(\frac{\lambda_1}{1-\lambda_{m+1}}x_1+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}x_m\Big)<\frac{\lambda_1}{1-\lambda_{m+1}}f(x_1)+\cdots+\frac{\lambda_m}{1-\lambda_{m+1}}f(x_m)$$

since

$$\frac{\lambda_1}{1-\lambda_{m+1}} + \cdots + \frac{\lambda_m}{1-\lambda_{m+1}} = 1.$$

Therefore, combining (3.3) and (3.4)

$$f(\lambda_1 x_1 + \cdots + \lambda_{m+1} x_{m+1}) < \lambda_1 f(x_1) + \cdots + \lambda_{m+1} f(x_{m+1}),$$

contradicting our hypotheses. Thus we must have $x_1 = \cdots = x_m$. From (3.5) we then conclude that $x_1 = \cdots = x_{m+1}$, as desired. ■

An interesting application of Jensen's inequality is the derivation of the so-called arithmetic/geometric mean inequalities. If $x_1, \dots, x_k \in \mathbb{R}_{>0}$, their ***arithmetic mean*** is

$$\frac{1}{k}(x_1 + \cdots + x_k)$$

and their ***geometric mean*** is

$$(x_1 \cdots x_k)^{1/k}.$$

We first state a result which relates generalisations of the arithmetic and geometric means.

3.1.35 Corollary (Weighted arithmetic/geometric mean inequality) *Let* $x_1, \dots, x_k \in \mathbb{R}_{\geq 0}$ *and suppose that* $\lambda_1, \dots, \lambda_k \in \mathbb{R}_{>0}$ *satisfy* $\sum_{j=1}^{k} \lambda_j = 1$. *Then*

$$x_1^{\lambda_1} \cdots x_k^{\lambda_k} \leq \lambda_1 x_1 + \cdots + \lambda_k x_k,$$

and equality holds if and only if $x_1 = \cdots = x_k$.

Proof Since the inequality obviously holds if any of $x_1, \dots, x_k$ are zero, let us suppose that these numbers are all positive. By Example 3.1.32–5, $-\log$ is convex. Thus Jensen's inequality gives

$$-\log(\lambda_1 x_1 + \cdots + \lambda_k x_k) \leq -\lambda_1 \log(x_1) - \cdots - \lambda_k \log(x_k) = -\log(x_1^{\lambda_1} \cdots x_k^{\lambda_k}).$$

Since $-\log$ is strictly monotonically decreasing by Proposition 3.6.6(ii), the result follows. Moreover, since $-\log$ is strictly convex by Proposition 3.2.30, the final assertion of the corollary follows from the final assertion of Theorem 3.1.34. ■

The corollary gives the following inequality as a special case.

3.1.36 Corollary (Arithmetic/geometric mean inequality) *If* $x_1, \dots, x_k \in \mathbb{R}_{\geq 0}$ *then*

$$(x_1 \cdots x_k)^{1/k} \leq \frac{x_1 + \cdots + x_k}{k},$$

and equality holds if and only if $x_1 = \cdots = x_k$.

Let us give some properties of convex functions. Further properties of convex function are give in Proposition 3.2.29

3.1.37 Proposition (Properties of convex functions I) *For an interval* $I \subseteq \mathbb{R}$ *and for a convex function* $f\colon I \to \mathbb{R}$*, the following statements hold:*

(i) if I *is open, then* f *is continuous;*

(ii) for any compact interval $K \subseteq \operatorname{int}(I)$*, there exists* $L \in \mathbb{R}_{>0}$ *such that*

$$|f(x_1) - f(x_2)| \le L|x_1 - x_2|, \qquad x_1, x_2 \in K.$$

Proof (ii) Let $K = [a, b] \subseteq \operatorname{int}(I)$ and let $a', b' \in I$ satisfy $a' < a$ and $b' > b$, this being possible since $K \subseteq \operatorname{int}(I)$. Now let $x_1, x_2 \in K$ and note that, by Lemma 3.1.33,

$$s_f(a', a) \le s_f(x_1, x_2) \le s_f(b, b')$$

since $a' < x_1$, $a \le x_2$, $x_1 \le b$, and $x_2 < b'$. Thus, taking $L = \max\{s_f(a', a), s_f(b, b')\}$, we have

$$-L \le \frac{f(x_2) - f(x_1)}{x_2 - x_1} \le L,$$

which gives the result.

(i) This follows from part (ii) easily. Indeed let $x \in I$ and let K be a compact subinterval of I such that $x \in \operatorname{int}(K)$, this being possible since I is open. If $\epsilon \in \mathbb{R}_{>0}$, let $\delta = \frac{\epsilon}{L}$. It then immediately follows that if $|x - y| < \delta$ then $|f(x) - f(y)| < \epsilon$. ■

Let us give some an example that illustrates that openness is necessary in the first part of the preceding result.

3.1.38 Example (A convex discontinuous function) Let $I = [0, 1]$ and define $f\colon [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, & x = 1, \\ 0, & x \in [0, 1). \end{cases}$$

If $x_1, x_2 \in [0, 1)$ and if $s \in [0, 1]$ then

$$0 = f((1 - s)x_1 + sx_2) = (1 - s)f(x_1) + sf(x_2).$$

If $x_1 \in [0, 1)$, if $x_2 = 1$, and if $s \in (0, 1)$ then

$$0 = f((1 - s)x_1 + sx_2) \le (1 - s)f(x_1) + sf(x_2) = s,$$

showing that f is convex as desired. Note that f is not continuous, but that its discontinuity is on the boundary, as must be the case since convex functions on open sets are continuous. •

Let us also present some operations that preserve convexity.

3.1.39 Proposition (Convexity and operations on functions) *For an interval* $I \subseteq \mathbb{R}$ *and for convex functions* $f, g\colon I \to \mathbb{R}$*, the following statements hold:*

(i) the function $I \ni x \mapsto \max\{f(x), g(x)\}$ *is convex;*

(ii) the function af *is convex if* $a \in \mathbb{R}_{\ge 0}$*;*

(iii) the function $f + g$ *is convex;*

(iv) *if* $J \subseteq \mathbb{R}$ *is an interval, if* f *takes values in* J*, and if* $\phi\colon J \to \mathbb{R}$ *is convex and monotonically increasing, then* $\phi \circ f$ *is convex;*

(v) *if* $x_0 \in I$ *is a local minimum for* f *(see Definition 3.2.15). then* x_0 *is a minimum for* f.

Proof (i) Let $x_1, x_2 \in I$ and let $s \in [0,1]$. Then, by directly applying the definition of convexity to f and g, we have

$$\begin{aligned}&\max\{f((1-s)x_1+sx_2), g((1-s)x_1+sx_2)\}\\ &\qquad\qquad\qquad \le (1-s)\max\{f(x_1),g(x_1)\} + s\max\{f(x_2),g(x_2)\}.\end{aligned}$$

(ii) This follows immediately from the definition of convexity.

(iii) For $x_1, x_2 \in I$ and for $s \in [0,1]$ we have

$$\begin{aligned}f((1-s)x_1+sx_2)+g((1-s)x_1+sx_2) &\le (1-s)f(x_1)+sf(x_2)+(1-s)g(x_1)+sg(x_2)\\ &= (1-s)(f(x_1)+g(x_1))+s(f(x_2+g(x_2)),\end{aligned}$$

by applying the definition of convexity to f and g.

(iv) For $x_1, x_2 \in I$ and for $s \in [0,1]$, convexity of f gives

$$f((1-s)x_1+sx_2) \le (1-s)f(x_1)+sf(x_2)$$

and so monotonicity of ϕ gives

$$\phi\circ f((1-s)x_1+sx_2) \le \phi((1-s)f(x_1)+sf(x_2)).$$

Now convexity of ϕ gives

$$\phi\circ f((1-s)x_1+sx_2) \le (1-s)\phi\circ f(x_1)+s\phi\circ f(x_2),$$

as desired.

(v) Suppose that x_0 is a local minimum for f, i.e., there is a neighbourhood $U \subseteq I$ of x_0 such that $f(x) \ge f(x_0)$ for all $x \in U$. Now let $x \in I$ and note that

$$s \mapsto (1-s)x_0 + sx$$

is continuous and $\lim_{s\to 0}(1-s)x_0+sx = x_0$. Therefore, there exists $s_0 \in (0,1]$ such that $(1-s)x_0+sx \in U$ for all $s \in (0,s_0)$. Thus

$$f(x_0) \le f((1-s)x_0+sx) \le (1-s)f(x_0)+sf(x)$$

for $s \in (0,s_0)$. Simplification gives $f(x_0) \le f(x)$ and so x_0 is a minimum for f. ■

3.1.7 Piecewise continuous functions

It is often of interest to consider functions that are not continuous, but which possess only jump discontinuities, and only “few” of these. In order to do so, it is convenient to introduce some notation. For and interval $I \subseteq \mathbb{R}$, a function $f\colon I \to \mathbb{R}$, and $x \in I$ define

$$f(x-) = \lim_{\epsilon\downarrow 0} f(x-\epsilon), \quad f(x+) = \lim_{\epsilon\downarrow 0} f(x+\epsilon),$$

allowing that these limits may not be defined (or even make sense if $x \in \mathrm{bd}(I)$).

We then have the following definition, recalling our notation concerning partitions of intervals given in and around Definition 2.5.7.

3.1.40 Definition (Piecewise continuous function) A function $f\colon [a,b] \to \mathbb{R}$ is ***piecewise continuous*** if there exists a partition $P = (I_1, \dots, I_k)$, with $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$, of $[a,b]$ with the following properties:

(i) $f|\,\mathrm{int}(I_j)$ is continuous for each $j \in \{1, \dots, k\}$;

(ii) for $j \in \{1, \dots, k-1\}$, the limits $f(x_j+)$ and $f(x_j-)$ exist;

(iii) the limits $f(a+)$ and $f(b-)$ exist. •

Let us give a couple of examples to illustrate some of the things that can happen with piecewise continuous functions.

3.1.41 Examples (Piecewise continuous functions)

1. Let $I = [-1,1]$ and define $f_1, f_2, f_3\colon I \to \mathbb{R}$ by

$$f_1(x) = \mathrm{sign}(x),$$
$$f_2(x) = \begin{cases} \mathrm{sign}(x), & x \neq 0, \\ 1, & x = 0, \end{cases}$$
$$f_2(x) = \begin{cases} \mathrm{sign}(x), & x \neq 0, \\ -1, & x = 0. \end{cases}$$

One readily verifies that all of these functions are piecewise continuous with a single discontinuity at $x = 0$. Note that the functions do not have the same value at the discontinuity. Indeed, the definition of piecewise continuity is unconcerned with the value of the function at discontinuities.

2. Let $I = [-1,1]$ and define $f\colon I \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

This function is, by definition, piecewise continuous with a single discontinuity at $x = 0$. This shows that the definition of piecewise continuity includes functions, not just with jump discontinuities, but with removable discontinuities. •

Exercises

3.1.1

Oftentimes, a continuity novice will think that the definition of continuity at x_0 of a function $f\colon I \to \mathbb{R}$ is as follows: for every $\epsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that if $|f(x) - f(x_0)| < \epsilon$ then $|x - x_0| < \delta$. Motivated by this, let us call a function ***fresh-from-high-school continuous*** if it has the preceding property at each point $x \in I$.

3.1.2 Answer the following two questions.

(a) Find an interval $I \subseteq \mathbb{R}$ and a function $f\colon I \to \mathbb{R}$ such that f is continuous but not fresh-from-high-school continuous.

(b) Find an interval $I \subseteq \mathbb{R}$ and a function $f\colon I \to \mathbb{R}$ such that f is fresh-from-high-school continuous but not continuous.

3.1.3 Let $I \subseteq \mathbb{R}$ be an interval and let $f, g\colon I \to \mathbb{R}$ be functions.

(a) Show that $D_{fg} \subseteq D_f \cup D_g$.

(b) Show that it is not generally true that $D_f \cap D_g \subseteq D_{fg}$.

(c) Suppose that f is bounded. Show that if $x \in (D_f \cap (I \setminus D_g)) \cap (I \setminus D_{fg})$, then $g(x) = 0$.***missing stuff***

3.1.4 Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}$ be a function. For $x \in I$ and $\delta \in \mathbb{R}_{>0}$ define

$$\omega_f(x, \delta) = \sup\{|f(x_1), f(x_2)| \mid x_1, x_2 \in \mathsf{B}(\delta, x) \cap I\}.$$

Show that, if $y \in \mathsf{B}(\delta, x)$, then $\omega_f(y, \frac{\delta}{2}) \leq \omega_f(x, \delta)$.

3.1.5 Recall from Theorem 3.1.24 that a continuous function defined on a compact interval is uniformly continuous. Show that this assertion is generally false if the interval is not compact.

3.1.6 Give an example of an interval $I \subseteq \mathbb{R}$ and a function $f\colon I \to \mathbb{R}$ that is locally bounded but not bounded.

3.1.7 Answer the following three questions.

(a) Find a bounded interval $I \subseteq \mathbb{R}$ and a function $f\colon I \to \mathbb{R}$ such that f is continuous but not bounded.

(b) Find a compact interval $I \subseteq \mathbb{R}$ and a function $f\colon I \to \mathbb{R}$ such that f is bounded but not continuous.

(c) Find a closed but unbounded interval $I \subseteq \mathbb{R}$ and a function $f\colon I \to \mathbb{R}$ such that f is continuous but not bounded.

3.1.8 Answer the following two questions.

(a) For $I = [0, 1)$ find a bounded, continuous function $f\colon I \to \mathbb{R}$ that does not attain its maximum on I.

(b) For $I = [0, \infty)$ find a bounded, continuous function $f\colon I \to \mathbb{R}$ that does not attain its maximum on I.

3.1.9 Explore your understanding of Theorem 3.1.3 and its Corollary 3.1.4 by doing the following.

(a) For the continuous function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$, verify Theorem 3.1.3 by (1) determining $f^{-1}(I)$ for a general open interval I and (2) showing that this is sufficient to ensure continuity.

Hint: *For the last part, consider using Proposition 2.5.6 and part (iv) of Proposition 1.3.5.*

(b) For the discontinuous function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \operatorname{sign}(x)$, verify Theorem 3.1.3 by (1) finding an open subset $U \subseteq \mathbb{R}$ for which $f^{-1}(U)$ is not open and (2) finding a sequence $(x_j)_{j \in \mathbb{Z}_{>0}}$ converging to $x_0 \in \mathbb{R}$ for which $(f(x_j))_{j \in \mathbb{Z}_{>0}}$ does not converge to $f(x_0)$.

3.1.10 Find a continuous function $f\colon I \to \mathbb{R}$ defined on some interval I and a sequence $(x_j)_{j \in \mathbb{Z}_{>0}}$ such that the sequence $(x_j)_{j \in \mathbb{Z}_{>0}}$ does not converge but the sequence $(f(x_j))_{j \in \mathbb{Z}_{>0}}$ does converge.

3.1.11 Let $I \subseteq \mathbb{R}$ be an interval and let $f, g\colon I \to \mathbb{R}$ be convex.

(a) Is it true that $x \mapsto \min\{f(x), g(x)\}$ is convex?

(b) Is it true that $f - g$ is convex?

3.1.12 Let $U \subseteq \mathbb{R}$ be open and suppose that $f\colon U \to \mathbb{R}$ is continuous and has the property that

$$\{x \in U \mid f(x) \neq 0\}$$

has measure zero. Show that $f(x) = 0$ for *all* $x \in U$.

Section 3.2

Differentiable $\mathbb{R}$-valued functions on $\mathbb{R}$

In this section we deal systematically with another topic with which most readers are at least somewhat familiar: differentiation. However, as with everything we do, we do this here is a manner that is likely more thorough and systematic than that seen by some readers. We do suppose that the reader has had that sort of course where one learns the derivatives of the standard functions, and learns to apply some of the standard rules of differentiation, such as we give in Section 3.2.3.

Do I need to read this section? If you are familiar with, or perhaps even if you only think you are familiar with, the meaning of "continuously differentiable," then you can probably forgo the details of this section. However, if you have not had the benefit of a rigorous calculus course, then the material here might at least be interesting. •

3.2.1 Definition of the derivative

The definition we give of the derivative is as usual, with the exception that, as we did when we talked about continuity, we allow functions to be defined on general intervals. In order to do this, we recall from Section 2.3.7 the notation $\lim_{x\to_I x_0} f(x)$.

3.2.1 Definition (Derivative and differentiable function) Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}$ be a function.

(i) The function f is ***differentiable at*** $\mathbf{x_0 \in I}$ if the limit

$$\lim_{x\to_I x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{3.6}$$

exists.

(ii) If the limit (3.6) exists, then it is denoted by $f'(x_0)$ and called the ***derivative*** of f at x_0.

(iii) If f is differentiable at each point $x \in I$, then f is ***differentiable***.

(iv) If f is differentiable and if the function $x \mapsto f'(x)$ is continuous, then f is ***continuously differentiable***, or of ***class*** $\mathbf{C^1}$. •

3.2.2 Notation (Alternative notation for derivative) In applications where $\mathbb{R}$-valued functions are clearly to be thought of as functions of "time," we shall sometimes write $\dot{f}$ rather than f' for the derivative.

Sometimes it is convenient to write the derivative using the convention $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}$. This notation for derivative suffers from the same problems as the notation "$f(x)$" to denote a function as discussed in Notation 1.3.2. That is to say, one cannot really use $\frac{\mathrm{d}f}{\mathrm{d}x}$ as a substitute for f', but only for $f'(x)$. Sometimes one can kludge one's way around this with something like $\frac{\mathrm{d}f}{\mathrm{d}x}\big|_{x=x_0}$ to specify the derivative at x_0. But this still leaves unresolved the matter of what is the rôle of "x" in the expression $\frac{\mathrm{d}f}{\mathrm{d}x}\big|_{x=x_0}$. For this reason, we will generally (but not exclusively) stick to f', or sometimes $\dot{f}$. For notation for the derivative for multivariable functions, we refer to Definition **??**. •

Let us consider some examples that illustrate the definition.

3.2.3 Examples (Derivative)

1. Take $I = \mathbb{R}$ and define $f\colon I \to \mathbb{R}$ by $f(x) = x^k$ for $k \in \mathbb{Z}_{>0}$. We claim that f is continuously differentiable, and that $f'(x) = kx^{k-1}$. To prove this we first note that
$$(x - x_0)(x^{k-1} + x^{k-1}x_0 + \cdots + xx_0^{k-2} + x_0^{k-1}) = x^k - x_0^k,$$
as can be directly verified. Then we compute
$$\begin{aligned}\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} &= \lim_{x\to x_0}\frac{x^k - x_0^k}{x-x_0}\\ &= \lim_{x\to x_0}(x^{k-1} + x^{k-1}x_0 + \cdots + xx_0^{k-2} + x_0^{k-1}) = kx_0^{k-1},\end{aligned}$$
as desired. Since f' is obviously continuous, we obtain that f is continuously differentiable, as desired.
2. Let $I = [0, 1]$ and define $f\colon I \to \mathbb{R}$ by
$$f(x) = \begin{cases} x, & x \neq 0,\\ 1, & x = 0.\end{cases}$$
From Example 1 we know that f is continuously differentiable at points in $(0, 1]$. We claim that f is not differentiable at $x = 0$. This will follow from Proposition 3.2.7 below, but let us show this here directly. We have
$$\lim_{x\to_I 0}\frac{f(x)-f(0)}{x-0} = \lim_{x\downarrow 0}\frac{x-1}{x} = -\infty.$$
Thus the limit does not exist, and so f is not differentiable at $x = 0$, albeit in a fairly stupid way.
3. Let $I = [0, 1]$ and define $f\colon I \to \mathbb{R}$ by $f(x) = \sqrt{x(1-x)}$. We claim that f is differentiable at points in $(0, 1)$, but is not differentiable at $x = 0$ or $x = 1$. Providing that one believes that the function $x \mapsto \sqrt{x}$ is differentiable on $\mathbb{R}_{>0}$ (see Section 3.6***missing stuff***), then the continuous differentiability of f on $(0, 1)$ follows from the results of Section 3.2.3. Moreover, the derivative of f at $x \in (0, 1)$ can be explicitly computed as
$$f'(x) = \frac{1-2x}{2\sqrt{x(1-x)}}.$$

To show that f is not differentiable at $x = 0$ we compute

$$\lim_{x\to_I 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\downarrow 0} \frac{\sqrt{1-x}}{\sqrt{x}} = \infty.$$

Similarly, at $x = 1$ we compute

$$\lim_{x\to_I 1} \frac{f(x) - f(1)}{x - 1} = \lim_{x\uparrow 1} \frac{-\sqrt{x}}{\sqrt{x-1}} = -\infty.$$

Since neither of these limits are elements of $\mathbb{R}$, it follows that f is not differentiable at $x = 0$ or $x = 1$.

4. Let $I = \mathbb{R}$ and define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

We first claim that f is differentiable. The differentiability of f at points $x \in \mathbb{R}\setminus\{0\}$ will follow from our results in Section 3.2.3 concerning differentiability, and algebraic operations along with composition. Indeed, using these rules for differentiation we compute that for $x \neq 0$ we have

$$f'(x) = 2x \sin \tfrac{1}{x} - \cos \tfrac{1}{x}.$$

Next let us prove that f is differentiable at $x = 0$ and that $f'(0) = 0$. We have

$$\lim_{x\to 0} \frac{f(x) - f(x)}{x - 0} = \lim_{x\to 0} x \sin \tfrac{1}{x}.$$

Now let $\epsilon \in \mathbb{R}_{>0}$. Then, for $\delta = \epsilon$ we have

$$\left| x \sin \tfrac{1}{x} - 0 \right| < \epsilon$$

since $\left|\sin \frac{1}{x}\right| \leq 1$. This shows that $f'(0) = 0$, as claimed. This shows that f is differentiable.

However, we claim that f is not *continuously* differentiable. Clearly there are no problems away from $x = 0$, again by the results of Section 3.2.3. But we note that f' is discontinuous at $x = 0$. Indeed, we note that f is the sum of two functions, one ($x \sin \frac{1}{x}$) of which goes to zero as x goes to zero, and the other ($-\cos \frac{1}{x}$) of which, when evaluated in any neighbourhood of $x = 0$, takes all possible values in the interval $[-1, 1]$. This means that in any sufficiently small neighbourhood of $x = 0$, the function f' will take all possible values in the interval $[-\frac{1}{2}, \frac{1}{2}]$. This precludes the limit $\lim_{x\to 0} f'(x)$ from existing, and so precludes f' from being continuous at $x = 0$ by Theorem 3.1.3. •

Let us give some intuition about the derivative. Given an interval I and functions $f, g\colon I \to \mathbb{R}$, we say that f and g are ***tangent*** at $x_0 \in \mathbb{R}$ if

$$\lim_{x\to_I x_0} \frac{f(x) - g(x)}{x - x_0} = 0.$$

In Figure 3.5 we depict the idea of two functions being tangent. Using this idea, we can give the following interpretation of the derivative.

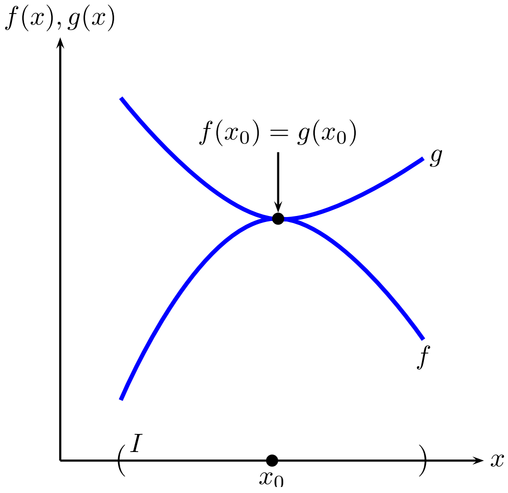


Figure 3.5 Functions that are tangent

3.2.4 Proposition (Derivative and linear approximation) *Let* $\mathrm{I} \subseteq \mathbb{R}$*, let* $\mathrm{x}_0 \in \mathrm{I}$*, and let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be a function. Then there exists at most one number* $\alpha \in \mathbb{R}$ *such that* f *is tangent at* x_0 *with the function* $\mathrm{x} \mapsto \mathrm{f}(\mathrm{x}_0) + \alpha(\mathrm{x} - \mathrm{x}_0)$*. Moreover, such a number* α *exists if and only if* f *is differentiable at* x_0*, in which case* $\alpha = \mathrm{f}'(\mathrm{x}_0)$*.*

Proof Suppose there are two such numbers α_1 and α_2. Thus

$$\lim_{x \to_I x_0} \frac{f(x) - (f(x_0) + \alpha_j(x - x_0))}{x - x_0} = 0, \qquad j \in \{1, 2\}, \tag{3.7}$$

We compute

$$\begin{aligned} |\alpha_1 - \alpha_2| &= \frac{|\alpha_1(x - x_0) - \alpha_2(x - x_0)|}{|x - x_0|} \\ &= \frac{|-f(x) + f(x_0) + \alpha_1(x - x_0) + f(x) - f(x_0) - \alpha_2(x - x_0)|}{|x - x_0|} \\ &\le \frac{|f(x) - f(x_0) - \alpha_1(x - x_0)|}{|x - x_0|} + \frac{|f(x) - f(x_0) - \alpha_2(x - x_0)|}{|x - x_0|}. \end{aligned}$$

Since α_1 and α_2 satisfy (3.7), as we let $x \to x_0$ the right-hand side goes to zero showing that $|\alpha_1 - \alpha_2| = 0$. This proves the first part of the result.

Next suppose that there exists $\alpha \in \mathbb{R}$ such that

$$\lim_{x \to_I x_0} \frac{f(x) - (f(x_0) + \alpha(x - x_0))}{x - x_0} = 0.$$

It then immediately follows that

$$\lim_{x \to_I x_0} \frac{f(x) - f(x_0)}{x - x_0} = \alpha.$$

Thus f is differentiable at x_0 with derivative equal to α. Conversely, if f is differentiable at x_0 then we have

$$\begin{aligned} f'(x_0) &= \lim_{x\to_I x_0} \frac{f(x)-f(x_0)}{x-x_0}, \\ \implies \quad \lim_{x\to_I x_0} &\frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0} = 0, \end{aligned}$$

which completes the proof. ■

The idea, then, is that the derivative serves, as we are taught in first-year calculus, as the best linear approximation to the function, since the function $x \mapsto f(x_0) + \alpha(x - x_0)$ is a linear function with slope α passing through $f(x_0)$.

We may also define derivatives of higher-order. Suppose that $f\colon I \to \mathbb{R}$ is differentiable, so that the function $f'\colon I \to \mathbb{R}$ can be defined. If the limit

$$\lim_{x\to_I x_0} \frac{f'(x)-f'(x_0)}{x-x_0}$$

exists, then we say that f is ***twice differentiable at* $\mathbf{x_0}$**. We denote the limit by $f''(x_0)$, and call it the ***second derivative*** of f at x_0. If f is differentiable at each point $x \in I$ then f is ***twice differentiable***. If additionally the map $x \mapsto f''(x)$ is continuous, then f is ***twice continuously differentiable***, or of ***class* $\mathbf{C^2}$**. Clearly this process can be continued inductively. Let us record the language coming from this iteration.

3.2.5 Definition (Higher-order derivatives) Let $I \subseteq \mathbb{R}$ be an interval, let $f\colon I \to \mathbb{R}$ be a function, let $r \in \mathbb{Z}_{>0}$, and suppose that f is $(r-1)$ times differentiable with g the corresponding $(r-1)$st derivative.

(i) The function f is **r *times differentiable at* $\mathbf{x_0 \in I}$** if the limit

$$\lim_{x\to_I x_0} \frac{g(x)-g(x_0)}{x-x_0} \tag{3.8}$$

exists.

(ii) If the limit (3.8) exists, then it is denoted by $f^{(r)}(x_0)$ and called the **r*th derivative*** of f at x_0.

(iii) If f is r times differentiable at each point $x \in I$, then f is **r *times differentiable***.

(iv) If f is r times differentiable and if the function $x \mapsto f^{(r)}(x)$ is continuous, then f is **r *times continuously differentiable***, or of ***class* $\mathbf{C^r}$**.

If f is of class C^r for each $r \in \mathbb{Z}_{>0}$, then f is ***infinitely differentiable***, or of ***class* $\mathbf{C^\infty}$**. •

3.2.6 Notation (Class $\mathbf{C^0}$) A continuous function will sometimes be said to be of ***class* $\mathbf{C^0}$**, in keeping with the language used for functions that are differentiable to some order. •

3.2.2 The derivative and continuity

In this section we simply do two things. We show that differentiable functions are continuous (Proposition 3.2.7), and we (dramatically) show that the converse of this is not true (Example 3.2.9).

3.2.7 Proposition (Differentiable functions are continuous) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is a function differentiable at* $\mathrm{x}_0 \in \mathrm{I}$*, then* f *is continuous at* x_0.

Proof Using Propositions 2.3.23 and 2.3.29 the limit

$$\lim_{x \to_I x_0} \Big(\frac{f(x) - f(x_0)}{x - x_0}\Big)(x - x_0)$$

exists, and is equal to the product of the limits

$$\lim_{x \to_I x_0} \frac{f(x) - f(x_0)}{x - x_0}, \qquad \lim_{x \to_I x_0} (x - x_0),$$

i.e., is equal to zero. We therefore can conclude that

$$\lim_{x \to_I x_0} (f(x) - f(x_0)) = 0,$$

and the result now follows from Theorem 3.1.3. ■

If the derivative is bounded, then there is more that one can say.

3.2.8 Proposition (Functions with bounded derivative are uniformly continuous) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is differentiable with* $\mathrm{f}'\colon \mathrm{I} \to \mathbb{R}$ *bounded, then* f *is uniformly continuous.*

Proof Let

$$M = \sup\{f'(t) \mid t \in I\}.$$

Then, for every $x, y \in I$, by the Mean Value Theorem, Theorem 3.2.19 below, there exists $z \in [x, y]$ such that

$$f(x) - f(y) = f'(z)(x - y) \quad \implies \quad |f(x) - f(y)| \leq M||x - y||.$$

Now let $\epsilon \in \mathbb{R}_{>0}$ and let $x \in I$. Define $\delta = \frac{\epsilon}{M}$ and note that if $y \in I$ satisfies $|x - y| < \delta$ then we have

$$|f(x) - f(y)| \leq M||x - y|| \leq \epsilon,$$

giving the desired uniform continuity. ■

Of course, it is not true that a continuous function is differentiable; we have an example of this as Example 3.2.3–3. However, things are much worse than that, as the following example indicates.

3.2.9 Example (A continuous but nowhere differentiable function) For $k \in \mathbb{Z}_{>0}$ define $g_k\colon \mathbb{R} \to \mathbb{R}$ as shown in Figure 3.6. Thus g_k is periodic with period $4 \cdot 2^{-2^k}$.[3] We then define

$$f(x) = \sum_{k=1}^{\infty} 2^{-k} g_k(x).$$

[3] We have not yet defined what is meant by a periodic function, although this is likely clear. In case it is not, a function $f\colon \mathbb{R} \to \mathbb{R}$ is ***periodic*** with period $T \in \mathbb{R}_{>0}$ if $f(x + T) = f(x)$ for every $x \in \mathbb{R}$. Periodic functions will be discussed in some detail in Section 8.1.6.

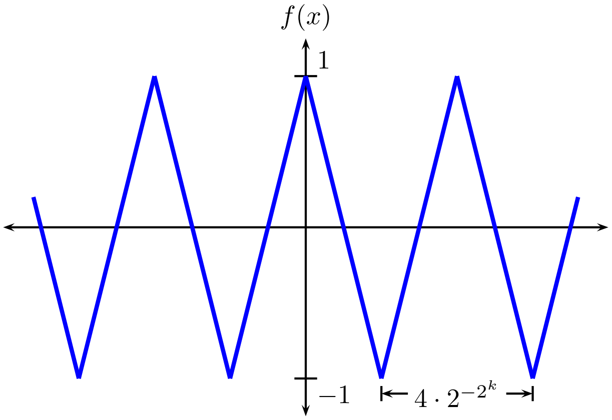


Figure 3.6 The function g_k

Since g_k is bounded in magnitude by 1, and since the sum $\sum_{k=1}^{\infty} 2^{-k}$ is absolutely convergent (Example 2.4.2–4), for each x the series defining f converges, and so f is well-defined. We claim that f is continuous but is nowhere differentiable.

It is easily shown by the Weierstrass M-test (see Theorem 3.5.15 below) that the series converges uniformly, and so defines a continuous function in the limit by Theorem 3.5.8. Thus f is continuous.

Now let us show that f is nowhere differentiable. Let $x \in \mathbb{R}$, $k \in \mathbb{Z}_{>0}$, and choose $h_k \in \mathbb{R}$ such that $|h| = 2^{-2^k}$ and such that x and $x + h_k$ lie on the line segment in the graph of g_k (this is possible since h_k is small enough, as is easily checked). Let us prove a few lemmata for this choice of x and h_k.

1 **Lemma** $\mathrm{g_l(x + h_k) = g(x)}$ *for* $\mathrm{l > k}$.

Proof This follows since g_l is periodic with period $4 \cdot 2^{-2^l}$, and so is therefore also periodic with period 2^{-2^k} since

$$\frac{4 \cdot 2^{-2^l}}{2^{-2^k}} = 4 \cdot 2^{-2^l - 2^k} \in \mathbb{Z}$$

for $l > k$. ▼

2 **Lemma** $\mathrm{|g_k(x + h_k) - g_k(x)| = 1}$.

Proof This follows from the fact that we have chosen h_k such that x and $x + h_k$ lie on the same line segment in the graph of g_k, and from the fact that $|h_k|$ is one-quarter the period of g_k (cf. Figure 3.6). ▼

3 **Lemma** $\mathrm{\left|\sum_{j=1}^{k-1} 2^{-j} g_j(x + h_k) - \sum_{j=1}^{k-1} 2^{-j} g_j(x)\right| \leq 2^k 2^{-2^{k-1}}}$.

Proof We note that if x and $x + h_k$ are on the same line segment in the graph of g_k, then they are also on the same line segment of the graph of g_j for $j \in \{1, \dots, k\}$.

Using this fact, along with the fact that the slope of the line segments of the function g_j have magnitude 2^{2^j}, we compute

$$\begin{aligned}\Big|\sum_{j=1}^{k-1} 2^{-j} g_j(x+h_k) - \sum_{j=1}^{k-1} 2^{-j} g_j(x)\Big| \\ &\leq (k-1)\max\{|2^{-j}g_j(x+h_k) - 2^{-j}g_j(x)| \mid j \in \{1,\dots,k\}\} \\ &= (k-1)2^{2^{k-1}}2^{-2^k} < 2^k 2^{-2^{k-1}}.\end{aligned}$$

The final inequality follows since $k-1 < 2^k$ for $k \geq 1$ and since $2^{2^{k-1}}2^{-2^k} = 2^{-2^{k-1}}$. ▼

Now we can assemble these lemmata to give the conclusion that f is not differentiable at x. Let $x \in \mathbb{R}$, let $\epsilon \in \mathbb{R}_{>0}$, choose $k \in \mathbb{Z}_{>0}$ such that $2^{-2^k} < \epsilon$, and choose h_k as above. We then have

$$\begin{aligned}\Big|\frac{f(x+h_k)-f(x)}{h_k}\Big| &= \Big|\frac{\sum_{j=1}^{\infty} 2^{-j}g_j(x+h_k) - \sum_{j=1}^{\infty} 2^{-j}g_j(x)}{h_k}\Big| \\ &= \Big|\frac{\sum_{j=1}^{k-1} 2^{-j}g_j(x+h_k) - \sum_{j=1}^{k-1} 2^{-j}g_j(x)}{h_k} + \frac{2^{-k}(g_k(x+h_k)-g_k(x))}{h_k}\Big| \\ &\geq 2^{-k}2^{2^k} - 2^k 2^{-2^{k-1}}.\end{aligned}$$

Since $\lim_{k\to\infty}(2^{-k}2^{2^k} - 2^k 2^{-2^{k-1}}) = \infty$, it follows that any neighbourhood of x will contain a point y for which $\frac{f(y)-f(x)}{y-x}$ will be as large in magnitude as desired. This precludes f from being differentiable at x. Now, since x was arbitrary in our construction, we have shown that f is nowhere differentiable as claimed.

In Figure 3.7 we plot the function

$$f_k(x) = \sum_{j=1}^{k} 2^{-j} g_j(x)$$

for $j \in \{1,2,3,4\}$. Note that, to the resolution discernible by the eye, there is no difference between f_3 and f_4. However, if we were to magnify the scale, we would see the effects that lead to the limit function not being differentiable. •

3.2.3 The derivative and operations on functions

In this section we provide the rules for using the derivative in conjunction with the natural algebraic operations on functions as described at the beginning of Section 3.1.3. Most readers will probably be familiar with these ideas, at least inasmuch as how to use them in practice.

3.2.10 Proposition (The derivative, and addition and multiplication) *Let* $I \subseteq \mathbb{R}$ *be an interval and let* $f, g\colon I \to \mathbb{R}$ *be functions differentiable at* $x_0 \in I$. *Then the following statements hold:*

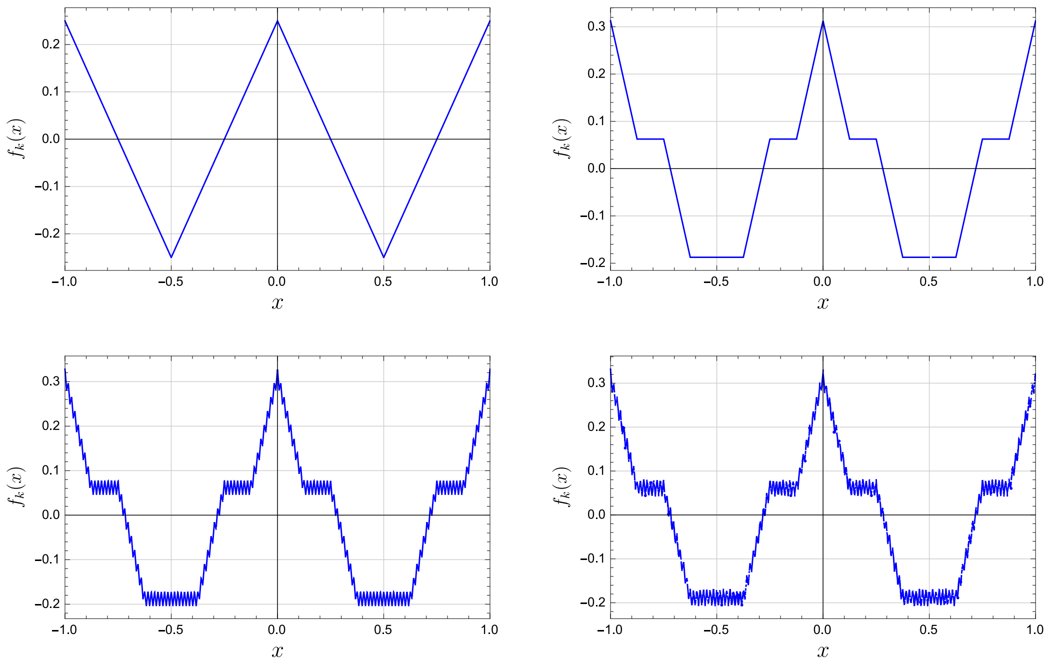


Figure 3.7 The first four partial sums for f

(i) $f + g$ *is differentiable at* x_0 *and* $(f + g)'(x_0) = f'(x_0) + g'(x_0)$*;*

(ii) fg *is differentiable at* x_0 *and* $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$ *(**product rule** or **Leibniz'**[4] **rule**);*

(iii) if additionally $g(x_0) \neq 0$*, then* $\frac{f}{g}$ *is differentiable at* x_0 *and*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2} \quad (\textbf{\textit{quotient rule}}).$$

Proof (i) We have

$$\frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0}.$$

Now we may apply Propositions 2.3.23 and 2.3.29 to deduce that

$$\lim_{x \to_I x_0} \frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} \\ = \lim_{x \to_I x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \to_I x_0} \frac{g(x) - g(x_0)}{x - x_0} = f'(x_0) + g'(x_0),$$

as desired.

[4] Gottfried Wilhelm von Leibniz (1646–1716) was born in Leipzig (then a part of Saxony), and was a lawyer, philosopher, and mathematician. His main mathematical contributions were to the development of calculus, where he had a well-publicised feud over priority with Newton, and algebra. His philosophical contributions, mainly in the area of logic, were also of some note.

(ii) Here we note that

$$\begin{aligned}\frac{(fg)(x)-(fg)(x_0)}{x-x_0} &= \frac{f(x)g(x)-f(x)g(x_0)+f(x)g(x_0)-f(x_0)g(x_0)}{x-x_0}\\ &= f(x)\frac{g(x)-g(x_0)}{x-x_0}+g(x_0)\frac{f(x)-f(x_0)}{x-x_0}.\end{aligned}$$

Since f is continuous at x_0 by Proposition 3.2.7, we may apply Propositions 2.3.23 and 2.3.29 to conclude that

$$\lim_{x\to_I x_0}\frac{(fg)(x)-(fg)(x_0)}{x-x_0}=f'(x_0)g(x_0)+f(x_0)g'(x_0),$$

just as claimed.

(iii) By using part (ii), it suffices to consider the case where f is defined by $f(x)=1$ (why?). Note that if $g(x_0)\neq 0$, then there is a neighbourhood of x_0 to which the restriction of g is nowhere zero. Thus, without loss of generality, we suppose that $g(x)\neq 0$ for all $x\in I$. But in this case we have

$$\lim_{x\to_I x_0}\frac{\frac{1}{g(x)}-\frac{1}{g(x_0)}}{x-x_0}=\lim_{x\to_I x_0}\frac{1}{g(x)g(x_0)}\frac{g(x_0)}{x-x_0}=-\frac{g'(x_0)}{g(x_0)^2},$$

giving the result in this case. We have used Propositions 2.3.23 and 2.3.29 as usual. ■

The following generalisation of the product rule will be occasionally useful.

3.2.11 Proposition (Higher-order product rule) *Let* $\mathrm{I}\subseteq\mathbb{R}$ *be an interval, let* $\mathrm{x}_0\in\mathrm{I}$*, let* $\mathrm{r}\in\mathbb{Z}_{>0}$*, and suppose that* $\mathrm{f},\mathrm{g}\colon\mathrm{I}\to\mathbb{R}$ *are of class* $\mathrm{C}^{\mathrm{r}-1}$ *and are* r*-times differentiable at* x_0*. Then* fg *is* r*-times differentiable at* x_0*, and*

$$(\mathrm{fg})^{(\mathrm{r})}(\mathrm{x}_0)=\sum_{\mathrm{j}=0}^{\mathrm{r}}\binom{\mathrm{r}}{\mathrm{j}}\mathrm{f}^{(\mathrm{j})}(\mathrm{x}_0)\mathrm{g}^{(\mathrm{r}-\mathrm{j})}(\mathrm{x}_0),$$

where

$$\binom{\mathrm{r}}{\mathrm{j}}=\frac{\mathrm{r}!}{\mathrm{j}!(\mathrm{r}-\mathrm{j})!}.$$

Proof The result is true for $r=1$ by Proposition 3.2.10. So suppose the result true for $k\in\{1,\dots,r\}$. We then have

$$\begin{aligned}\frac{(fg)^{(r)}(x)-(fg)^{(r)}(x_0)}{x-x_0} &= \frac{\sum_{j=0}^{r}\binom{r}{j}f^{(j)}(x)g^{(r-j)}(x)-\sum_{j=0}^{r}\binom{r}{j}f^{(j)}(x_0)g^{(r-j)}(x_0)}{x-x_0}\\ &= \sum_{j=0}^{r}\binom{r}{j}\frac{f^{(j)}(x)g^{(r-j)}(x)-f^{(j)}(x_0)g^{(r-j)}(x_0)}{x-x_0}.\end{aligned}$$

Now we note that

$$\lim_{x\to_I x_0}\frac{f^{(j)}(x)g^{(r-j)}(x)-f^{(j)}(x_0)g^{(r-j)}(x_0)}{x-x_0}=f^{(j+1)}(x_0)g^{(r-j)}(x_0)+f^{(j)}(x_0)g^{(r-j+1)}(x_0).$$

Therefore,

$$\begin{aligned}
&\lim_{x\to_I x_0}\frac{(fg)^{(r)}(x)-(fg)^{(r)}(x_0)}{x-x_0}\\
&\quad=\sum_{j=0}^{r}\binom{r}{j}\Big(f^{(j+1)}(x_0)g^{(r-j)}(x_0)+f^{(j)}(x_0)g^{(r-j+1)}(x_0)\Big)\\
&\quad=f(x_0)g^{(r+1)}(x_0)+\sum_{j=0}^{r}\binom{r}{j}f^{(j+1)}(x_0)g^{(r-j)}(x_0)+\sum_{j=1}^{r}\binom{r}{j}f^{(j)}(x_0)g^{(r-j+1)}(x_0)\\
&\quad=f(x_0)g^{(r+1)}(x_0)+\sum_{j=1}^{r+1}\binom{r}{j-1}f^{(j)}(x_0)g^{(r-j+1)}(x_0)\\
&\qquad+\sum_{j=1}^{r}\binom{r}{j}f^{(j)}(x_0)g^{(r-j+1)}(x_0)\\
&\quad=f^{(r+1)}(x_0)g(x_0)+f(x_0)g^{(r+1)}(x_0)\\
&\qquad+\sum_{j=1}^{r}\left(\binom{r}{j}+\binom{r}{j-1}\right)f^{(j)}(x_0)g^{(r-j+1)}(x_0)\\
&\quad=f^{(r+1)}(x_0)g(x_0)+f(x_0)g^{(r+1)}(x_0)+\sum_{j=1}^{r}\binom{r+1}{j}f^{(j)}(x_0)g^{(r-j+1)}(x_0)\\
&\quad=\sum_{j=0}^{r+1}\binom{r+1}{j}f^{(j)}(x_0)g^{(r-j)}(x_0).
\end{aligned}$$

In the penultimate step we have used ***Pascal's***[5] ***Rule*** which states that

$$\binom{r}{j}+\binom{r}{j-1}=\binom{r+1}{j}.$$

We leave the direct proof of this fact to the reader. ■

The preceding two results had to do with differentiability at a point. For convenience, let us record the corresponding results when we consider the derivative, not just at a point, but on the entire interval.

3.2.12 Proposition (Class C^r, and addition and multiplication) *Let* $I\subseteq\mathbb{R}$ *be an interval and let* $f,g\colon I\to\mathbb{R}$ *be functions of class* C^r*. Then the following statements hold:*

(i) $f+g$ *is of class* C^r*;*

(ii) fg *is of class* C^r*;*

(iii) if additionally $g(x)\neq 0$ *for all* $x\in I$*, then* $\frac{f}{g}$ *is of class* C^r*.*

Proof This follows directly from Propositions 3.2.10 and 3.2.11, along with the fact, following from Proposition 3.1.15, that the expressions for the derivatives of sums, products, and quotients are continuous, as they are themselves sums, products, and quotients. ■

[5] Blaise Pascal (1623–1662) was a French mathematician and philosopher. Much of his mathematical work was on analytic geometry and probability theory.

The following rule for differentiating the composition of functions is one of the more useful of the rules concerning the behaviour of the derivative.

3.2.13 Theorem (Chain Rule) *Let* $\mathrm{I}, \mathrm{J} \subseteq \mathbb{R}$ *be intervals and let* $\mathrm{f}\colon \mathrm{I} \to \mathrm{J}$ *and* $\mathrm{g}\colon \mathrm{J} \to \mathbb{R}$ *be functions for which* f *is differentiable at* $\mathrm{x}_0 \in \mathrm{I}$ *and* g *is differentiable at* $\mathrm{f}(\mathrm{x}_0) \in \mathrm{J}$*. Then* $\mathrm{g} \circ \mathrm{f}$ *is differentiable at* x_0*, and* $(\mathrm{g} \circ \mathrm{f})'(\mathrm{x}_0) = \mathrm{g}'(\mathrm{f}(\mathrm{x}_0))\mathrm{f}'(\mathrm{x}_0)$*.*

Proof Let us define $h\colon J \to \mathbb{R}$ by

$$h(y) = \begin{cases} \frac{g(y)-g(f(x_0))}{y-f(x_0)}, & g(y) \neq g(f(x_0)), \\ g'(f(x_0)), & g(y) = g(f(x_0)). \end{cases}$$

We have

$$\frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = \frac{(g \circ f)(x) - (g \circ f)(x_0)}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0} = h(f(x)) \frac{f(x) - f(x_0)}{x - x_0},$$

provided that $f(x) \neq f(x_0)$. On the other hand, if $f(x) = f(x_0)$, we immediately have

$$\frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = h(f(x)) \frac{f(x) - f(x_0)}{x - x_0}$$

since both sides of this equation are zero. Thus we simply have

$$\frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = h(f(x)) \frac{f(x) - f(x_0)}{x - x_0}$$

for all $x \in I$. Note that h is continuous at $f(x_0)$ by Theorem 3.1.3 since

$$\lim_{y \to_I f(x_0)} h(y) = g'(x_0) = h(x_0),$$

using the fact that g is differentiable at x_0. Now we can use Propositions 2.3.23 and 2.3.29 to ascertain that

$$\lim_{x \to_I x_0} \frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = \lim_{x \to_I x_0} h(f(x)) \frac{f(x) - f(x_0)}{x - x_0} = g'(f(x_0))f'(x_0),$$

as desired. ■

The derivative behaves as one would expect when restricting a differentiable function.

3.2.14 Proposition (The derivative and restriction) *If* $\mathrm{I}, \mathrm{J} \subseteq \mathbb{R}$ *are intervals for which* $\mathrm{J} \subseteq \mathrm{I}$*, and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is differentiable at* $\mathrm{x}_0 \in \mathrm{J} \subseteq \mathrm{I}$*, then* $\mathrm{f}|\mathrm{J}$ *is differentiable at* x_0*.*

Proof This follows since if the limit

$$\lim_{x \to_I x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists, then so too does the limit

$$\lim_{x \to_J x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

provided that $J \subseteq I$. ■

missing stuff

3.2.4 The derivative and function behaviour

From the behaviour of the derivative of a function, it is often possible to deduce some important features of the function itself. One of the most important of these concerns maxima and minima of a function. Let us define these concepts precisely.

3.2.15 Definition (Local maximum and local minimum) Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}$ be a function. A point $x_0 \in I$ is a:

(i) ***local maximum*** if there exists a neighbourhood U of x_0 such that $f(x) \leq f(x_0)$ for every $x \in U$;

(ii) ***strict local maximum*** if there exists a neighbourhood U of x_0 such that $f(x) < f(x_0)$ for every $x \in U \setminus \{x_0\}$;

(iii) ***local minimum*** if there exists a neighbourhood U of x_0 such that $f(x) \geq f(x_0)$ for every $x \in U$;

(iv) ***strict local minimum*** if there exists a neighbourhood U of x_0 such that $f(x) > f(x_0)$ for every $x \in U \setminus \{x_0\}$. •

Now we have the standard result that relates derivatives to maxima and minima.

3.2.16 Theorem (Derivatives, and maxima and minima) *For* $\mathrm{I} \subseteq \mathbb{R}$ *an interval,* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *a function, and* $\mathrm{x}_0 \in \mathrm{int}(\mathrm{I})$*, the following statements hold:*

(i) if f *is differentiable at* x_0 *and if* x_0 *is a local maximum or a local minimum for* f*, then* $\mathrm{f}'(\mathrm{x}_0) = 0$*;*

(ii) if f *is twice differentiable at* x_0*, and if* x_0 *is a local maximum (resp. local minimum) for* f*, then* $\mathrm{f}''(\mathrm{x}_0) \leq 0$ *(resp.* $\mathrm{f}''(\mathrm{x}_0) \geq 0$*);*

(iii) if f *is twice differentiable at* x_0*, and if* $\mathrm{f}'(\mathrm{x}_0) = 0$ *and* $\mathrm{f}''(\mathrm{x}_0) \in \mathbb{R}_{<0}$ *(resp.* $\mathrm{f}''(\mathrm{x}_0) \in \mathbb{R}_{>0}$*), then* x_0 *is a strict local maximum (resp. strict local minimum) for* f*.*

Proof (i) We will prove the case where x_0 is a local minimum, since the case of a local maximum is similar. If x_0 is a local minimum, then there exists $\epsilon \in \mathbb{R}_{>0}$ such that $f(x) \geq f(x_0)$ for all $x \in \mathsf{B}(\epsilon, x_0)$. Therefore, $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$ for $x \geq x_0$ and $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$ for $x \leq x_0$. Since the limit $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ exists, it must be equal to both limits

$$\lim_{x\downarrow x_0} \frac{f(x)-f(x_0)}{x-x_0}, \quad \lim_{x\uparrow x_0} \frac{f(x)-f(x_0)}{x-x_0}.$$

However, since the left limit is nonnegative and the right limit is nonpositive, we conclude that $f'(x_0) = 0$.

(ii) We shall show that if f is twice differentiable at x_0 and $f''(x_0)$ is not less than or equal to zero, then x_0 is not a local maximum. The statement concerning the local minimum is argued in the same way. Now, if f is twice differentiable at x_0, and if $f''(x_0) \in \mathbb{R}_{>0}$, then x_0 is a local minimum by part (iii), which prohibits it from being a local maximum.

(iii) We consider the case where $f''(x_0) \in \mathbb{R}_{>0}$, since the other case follows in the same manner. Choose $\epsilon \in \mathbb{R}_{>0}$ such that, for $x \in \mathsf{B}(\epsilon, x_0)$,

$$\left|\frac{f'(x)-f'(x_0)}{x-x_0} - f''(x_0)\right| < \tfrac{1}{2} f''(x_0),$$

this being possible since $f''(x_0) > 0$ and since f is twice differentiable at x_0. Since $f''(x_0) > 0$ it follows that, for $x \in \mathsf{B}(\epsilon, x_0)$,

$$\frac{f'(x) - f'(x_0)}{x - x_0} > 0,$$

from which we conclude that $f'(x) > 0$ for $x \in (x_0, x_0 + \epsilon)$ and that $f'(x) < 0$ for $x \in (x_0 - \epsilon, x_0)$. Now we prove a technical lemma.

1 Lemma *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an open interval, let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be a continuous function that is differentiable, except possibly at* $\mathrm{x}_0 \in \mathrm{I}$. *If* $\mathrm{f}'(\mathrm{x}) > 0$ *for every* $\mathrm{x} > \mathrm{x}_0$ *and if* $\mathrm{f}'(\mathrm{x}) < 0$ *for every* $\mathrm{x} < \mathrm{x}_0$, *then* x_0 *is a strict local minimum for* f.

Proof We will use the Mean Value Theorem (Theorem 3.2.19) which we prove below. Note that our proof of the Mean Value Theorem depends on part (i) of the present theorem, but not on part that we are now proving. Let $x \in I \setminus \{x_0\}$. We have two cases.

1. $x > x_0$: By the Mean Value Theorem there exists $a \in (x, x_0)$ such that $f(x) - f(x_0) = (x - x_0)f'(a)$. Since $f'(a) > 0$ it then follows that $f(x) > f(x_0)$.
2. $x < x_0$: A similar argument as in the previous case again gives $f(x) > f(x_0)$.

Combining these conclusions, we see that $f(x) > f(x_0)$ for all $x \in I$, and so x_0 is a strict local maximum for f. ▼

The lemma now immediately applies to the restriction of f to $\mathsf{B}(\epsilon, x_0)$, and so gives the result. ■

Let us give some examples that illustrate the value and limitations of the preceding result.

3.2.17 Examples (Derivatives, and maxima and minima)

1. Let $I = \mathbb{R}$ and define $f\colon I \to \mathbb{R}$ by $f(x) = x^2$. Note that f is infinitely differentiable, so Theorem 3.2.16 can be applied freely. We compute $f'(x) = 2x$, and so $f'(x) = 0$ if and only if $x = 0$. Therefore, the only local maxima and local minima must occur at $x = 0$. To check whether a local maxima, a local minima, or neither exists at $x = 0$, we compute the second derivative which is $f''(x) = 2$. This is positive at $x = 0$ (and indeed everywhere), so we may conclude that $x = 0$ is a strict local maximum for f from part (iii) of the theorem.

   Applying the same computations to $g(x) = -x^2$ shows that $x = 0$ is a strict local maximum for g.
2. Let $I = \mathbb{R}$ and define $f\colon I \to \mathbb{R}$ by $f(x) = x^3$. We compute $f'(x) = 3x^2$, from which we ascertain that all maxima and minima must occur, if at all, at $x = 0$. However, since $f''(x) = 6x$, $f''(0) = 0$, and we cannot conclude from Theorem 3.2.16 whether there is a local maximum, a local minimum, or neither at $x = 0$. In fact, one can see "by hand" that $x = 0$ is neither a local maximum nor a local minimum for f.

   The same arguments apply to the functions $g(x) = x^4$ and $h(x) = -x^4$ to show that when the second derivative vanishes, it is possible to have all possibilities—a local maximum, a local minimum, or neither—at a point where both f' and f'' are zero.

3. Let $I = [-1,1]$ and define $f\colon I \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1-x, & x \in [0,1], \\ 1+x, & x \in [-1,0). \end{cases}$$

"By hand," one can check that f has a strict local maximum at $x = 0$, and strict local minima at $x = -1$ and $x = 1$. However, we can detect none of these using Theorem 3.2.16. Indeed, the local minima at $x = -1$ and $x = 1$ occur at the boundary of I, and so the hypotheses of the theorem do not apply. This, indeed, is why we demand that x_0 lie in $\operatorname{int}(I)$ in the theorem statement. For the local maximum at $x = 0$, the theorem does not apply since f is not differentiable at $x = 0$. However, we do note that Lemma 1 (with modifications to the signs of the derivative in the hypotheses, and changing "minimum" to "maximum" in the conclusions) in the proof of the theorem *does* apply, since f is differentiable at points in $(-1,0)$ and $(0,1)$, and for $x > 0$ we have $f'(x) < 0$ and for $x < 0$ we have $f'(x) > 0$. The lemma then allows us to conclude that f has a strict local maximum at $x = 0$. •

Next let us prove a simple result that, while not always of great value itself, leads to the important Mean Value Theorem below.

3.2.18 Theorem (Rolle's[6] Theorem) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval, let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be continuous, and suppose that for* $\mathrm{a}, \mathrm{b} \in \mathrm{I}$ *it holds that* $\mathrm{f}|(\mathrm{a},\mathrm{b})$ *is differentiable and that* $\mathrm{f}(\mathrm{a}) = \mathrm{f}(\mathrm{b})$*. Then there exists* $\mathrm{c} \in (\mathrm{a},\mathrm{b})$ *such that* $\mathrm{f}'(\mathrm{c}) = 0$*.*

Proof Since $f|[a,b]$ is continuous, by Theorem 3.1.23 there exists $x_1, x_2 \in [a,b]$ such that $\operatorname{image}(f|[a,b]) = [f(x_1), f(x_2)]$. We have three cases to consider.

1. $x_1, x_2 \in \operatorname{bd}([a,b])$: In this case it holds that f is constant since $f(a) = f(b)$. Thus the conclusions of the theorem hold for any $c \in (a,b)$.
2. $x_1 \in \operatorname{int}([a,b])$: In this case, f has a local minimum at x_1, and so by Theorem 3.2.16(i) we conclude that $f'(x_1) = 0$.
3. $x_2 \in \operatorname{int}([a,b])$: In this case, f has a local maximum at x_2, and so by Theorem 3.2.16(i) we conclude that $f'(x_2) = 0$. ■

Rolle's Theorem has the following generalisation, which is often quite useful, since it establishes links between the values of a function and the values of its derivative.

3.2.19 Theorem (Mean Value Theorem) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval, let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be continuous, and suppose that for* $\mathrm{a}, \mathrm{b} \in \mathrm{I}$ *it holds that* $\mathrm{f}|(\mathrm{a},\mathrm{b})$ *is differentiable. Then there exists* $\mathrm{c} \in (\mathrm{a},\mathrm{b})$ *such that*

$$\mathrm{f}'(\mathrm{c}) = \frac{\mathrm{f}(\mathrm{b}) - \mathrm{f}(\mathrm{a})}{\mathrm{b} - \mathrm{a}}.$$

Proof Define $g\colon I \to \mathbb{R}$ by

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

[6] Michel Rolle (1652–1719) was a French mathematician whose primary contributions were to algebra.

Using the results of Section 3.2.3 we conclude that g is continuous and differentiable on (a, b). Moreover, direct substitution shows that $g(b) = g(a)$. Thus Rolle's Theorem allows us to conclude that there exists $c \in (a, b)$ such that $g'(c) = 0$. However, another direct substitution shows that $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$. ■

In Figure 3.8 we give the intuition for Rolle's Theorem, the Mean Value Theorem, and the relationship between the two results.

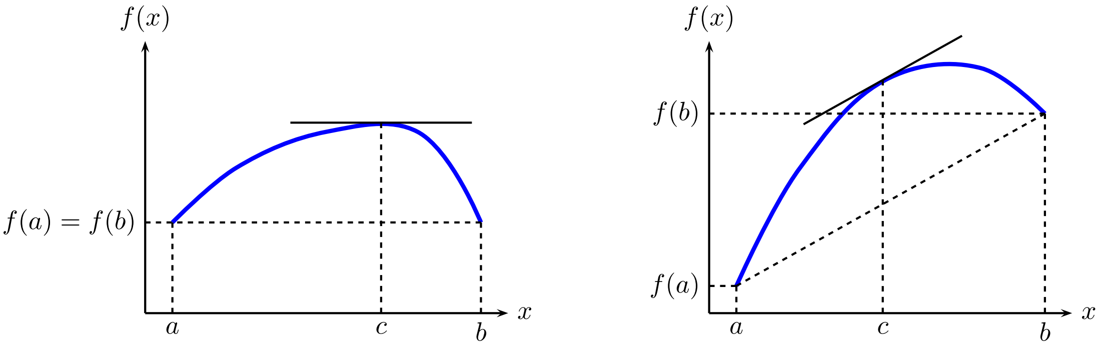


Figure 3.8 Illustration of Rolle's Theorem (left) and the Mean Value Theorem (right)

Another version of the Mean Value Theorem relates the values of two functions with the values of their derivatives.

3.2.20 Theorem (Cauchy's Mean Value Theorem) *Let* $I \subseteq \mathbb{R}$ *be an interval and let* $f, g \colon I \rightarrow \mathbb{R}$ *be continuous, and suppose that for* $a, b \in I$ *it holds that* $f|(a, b)$ *and* $g|(a, b)$ *are differentiable, and that* $g'(x) \neq 0$ *for each* $x \in (a, b)$*. Then there exists* $c \in (a, b)$ *such that*

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof Note that $g(b) \neq g(a)$ by Rolle's Theorem, since $g'(x) \neq 0$ for $x \in \operatorname{int}(a, b)$. Let

$$\alpha = \frac{f(b) - f(a)}{g(b) - g(a)}$$

and define $h \colon I \rightarrow \mathbb{R}$ by $h(x) = f(x) - \alpha g(x)$. Using the results of Section 3.2.3, one verifies that h is continuous on I and differentiable on (a, b). Moreover, one can also verify that $h(a) = h(b)$. Thus Rolle's Theorem implies the existence of $c \in (a, b)$ for which $h'(c) = 0$. A simple computation verifies that $h'(c) = 0$ is equivalent to the conclusion of the theorem. ■

We conclude this section with the useful L'Hôpital's Rule. This rule for finding limits is sufficiently useful that we state and prove it here in an unusual level of generality.

3.2.21 Theorem (L'Hôpital's[7] Rule) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval, let* $\mathrm{x}_0 \in \mathbb{R}$*, and let* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *be differentiable functions with* $\mathrm{g}'(\mathrm{x}) \neq 0$ *for all* $\mathrm{x} \in \mathrm{I} - \{\mathrm{x}_0\}$*. Then the following statements hold.*

(i) *Suppose that* x_0 *is an open right endpoint for* I *and suppose that either*

(a) $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \mathrm{f}(\mathrm{x}) = 0$ *and* $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \mathrm{g}(\mathrm{x}) = 0$ *or*

(b) $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \mathrm{f}(\mathrm{x}) = \infty$ *and* $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \mathrm{g}(\mathrm{x}) = \infty$*,*

and suppose that $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \frac{\mathrm{f}'(\mathrm{x})}{\mathrm{g}'(\mathrm{x})} = \mathrm{s}_0 \in \overline{\mathbb{R}}$*. Then* $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \frac{\mathrm{f}(\mathrm{x})}{\mathrm{g}(\mathrm{x})} = \mathrm{s}_0$.

(ii) *Suppose that* x_0 *is an left right endpoint for* I *and suppose that either*

(a) $\lim_{\mathrm{x}\downarrow\mathrm{x}_0} \mathrm{f}(\mathrm{x}) = 0$ *and* $\lim_{\mathrm{x}\downarrow\mathrm{x}_0} \mathrm{g}(\mathrm{x}) = 0$ *or*

(b) $\lim_{\mathrm{x}\uparrow\mathrm{x}_0} \mathrm{f}(\mathrm{x}) = \infty$ *and* $\lim_{\mathrm{x}\downarrow\mathrm{x}_0} \mathrm{g}(\mathrm{x}) = \infty$*,*

and suppose that $\lim_{\mathrm{x}\downarrow\mathrm{x}_0} \frac{\mathrm{f}'(\mathrm{x})}{\mathrm{g}'(\mathrm{x})} = \mathrm{s}_0 \in \overline{\mathbb{R}}$*. Then* $\lim_{\mathrm{x}\downarrow\mathrm{x}_0} \frac{\mathrm{f}(\mathrm{x})}{\mathrm{g}(\mathrm{x})} = \mathrm{s}_0$.

(iii) *Suppose that* $\mathrm{x}_0 \in \mathrm{int}(\mathrm{I})$ *and suppose that either*

(a) $\lim_{\mathrm{x}\to\mathrm{x}_0} \mathrm{f}(\mathrm{x}) = 0$ *and* $\lim_{\mathrm{x}\to\mathrm{x}_0} \mathrm{g}(\mathrm{x}) = 0$ *or*

(b) $\lim_{\mathrm{x}\to\mathrm{x}_0} \mathrm{f}(\mathrm{x}) = \infty$ *and* $\lim_{\mathrm{x}\to\mathrm{x}_0} \mathrm{g}(\mathrm{x}) = \infty$*,*

and suppose that $\lim_{\mathrm{x}\to\mathrm{x}_0} \frac{\mathrm{f}'(\mathrm{x})}{\mathrm{g}'(\mathrm{x})} = \mathrm{s}_0 \in \overline{\mathbb{R}}$*. Then* $\lim_{\mathrm{x}\to\mathrm{x}_0} \frac{\mathrm{f}(\mathrm{x})}{\mathrm{g}(\mathrm{x})} = \mathrm{s}_0$.

The following two statements which are independent of x_0 *(thus we ask that* $\mathrm{g}'(\mathrm{x}) \neq 0$ *for all* $\mathrm{x} \in \mathrm{I}$*) also hold.*

(iv) *Suppose that* I *is unbounded on the right and suppose that either*

(a) $\lim_{\mathrm{x}\to\infty} \mathrm{f}(\mathrm{x}) = 0$ *and* $\lim_{\mathrm{x}\to\infty} \mathrm{g}(\mathrm{x}) = 0$ *or*

(b) $\lim_{\mathrm{x}\to\infty} \mathrm{f}(\mathrm{x}) = \infty$ *and* $\lim_{\mathrm{x}\to\infty} \mathrm{g}(\mathrm{x}) = \infty$*,*

and suppose that $\lim_{\mathrm{x}\to\infty} \frac{\mathrm{f}'(\mathrm{x})}{\mathrm{g}'(\mathrm{x})} = \mathrm{s}_0 \in \overline{\mathbb{R}}$*. Then* $\lim_{\mathrm{x}\to\infty} \frac{\mathrm{f}(\mathrm{x})}{\mathrm{g}(\mathrm{x})} = \mathrm{s}_0$.

(v) *Suppose that* I *is unbounded on the left and suppose that either*

(a) $\lim_{\mathrm{x}\to-\infty} \mathrm{f}(\mathrm{x}) = 0$ *and* $\lim_{\mathrm{x}\to-\infty} \mathrm{g}(\mathrm{x}) = 0$ *or*

(b) $\lim_{\mathrm{x}\to-\infty} \mathrm{f}(\mathrm{x}) = \infty$ *and* $\lim_{\mathrm{x}\to-\infty} \mathrm{g}(\mathrm{x}) = \infty$*,*

and suppose that $\lim_{\mathrm{x}\to-\infty} \frac{\mathrm{f}'(\mathrm{x})}{\mathrm{g}'(\mathrm{x})} = \mathrm{s}_0 \in \overline{\mathbb{R}}$*. Then* $\lim_{\mathrm{x}\to-\infty} \frac{\mathrm{f}(\mathrm{x})}{\mathrm{g}(\mathrm{x})} = \mathrm{s}_0$.

Proof (i) First suppose that $\lim_{x\uparrow x_0} f(x) = 0$ and $\lim_{x\uparrow x_0} g(x) = 0$ and that $s_0 \in \mathbb{R}$. We may then extend f and g to be defined at x_0 by taking their values at x_0 to be zero, and the resulting function will be continuous by Theorem 3.1.3. We may now apply Cauchy's Mean Value Theorem to assert that for $x \in I$ there exists $c_x \in (x, x_0)$ such that

$$\frac{f'(c_x)}{g'(c_x)} = \frac{f(x_0) - f(x)}{g(x_0) - g(x)} = \frac{f(x)}{g(x)}.$$

Now let $\epsilon \in \mathbb{R}_{>0}$ and choose $\delta \in \mathbb{R}_{>0}$ such that $\left|\frac{f'(x)}{g'(x)} - s_0\right| < \epsilon$ for $x \in \mathsf{B}(\delta, x_0) \cap I$. Then, for $x \in \mathsf{B}(\delta, x_0) \cap I$ we have

$$\left|\frac{f(x)}{g(x)} - s_0\right| = \left|\frac{f'(c_x)}{g'(c_x)} - s_0\right| < \epsilon$$

[7] Guillaume François Antoine Marquis de L'Hôpital (1661–1704) was one of the early developers of calculus.

since $c_x \in \mathsf{B}(\delta, x_0) \cap I$. This shows that $\lim_{x\uparrow x_0} \frac{f(x)}{g(x)} = s_0$, as claimed.

Now suppose that $\lim_{x\uparrow x_0} f(x) = \infty$ and $\lim_{x\uparrow x_0} g(x) = \infty$ and that $s_0 \in \mathbb{R}$. Let $\epsilon \in \mathbb{R}_{>0}$ and choose $\delta_1 \in \mathbb{R}_{>0}$ such that $\left|\frac{f'(x)}{g'(x)} - s_0\right| < \frac{\epsilon}{2(1+|s_0|)}$ for $x \in \mathsf{B}(\delta_1, x_0) \cap I$. For $x \in \mathsf{B}(\delta_1, x_0) \cap I$, by Cauchy's Mean Value Theorem there exists $c_x \in \mathsf{B}(\delta_1, x_0) \cap I$ such that

$$\frac{f'(c_x)}{g'(c_x)} = \frac{f(x) - f(x-\delta_1)}{g(x) - g(x-\delta_1)}.$$

Therefore,

$$\left|\frac{f(x) - f(x-\delta_1)}{g(x) - g(x-\delta_1)} - s_0\right| < \frac{\epsilon}{2(1+|s_0|)}$$

for $x \in \mathsf{B}(\delta, x_0) \cap I$. Now define

$$h(x) = \frac{1 - \frac{f(x-\delta_1)}{f(x)}}{1 - \frac{g(x-\delta_1)}{g(x)}}$$

and note that

$$\frac{f(x) - f(x-\delta_1)}{g(x) - g(x-\delta_1)} = h(x)\frac{f(x)}{g(x)}.$$

Therefore we have

$$\left|h(x)\frac{f(x)}{g(x)} - s_0\right| < \frac{\epsilon}{2(1+|s_0|)}$$

for $x \in \mathsf{B}(\delta_1, x_0) \cap I$. Note also that $\lim_{x\uparrow x_0} h(x) = 1$. Thus we can choose $\delta_2 \in \mathbb{R}_{>0}$ such that $|h(x) - 1| < \frac{\epsilon}{2(1+|s_0|)}$ and $h(x) > \frac{1}{2}$ for $x \in \mathsf{B}(\delta_2, x_0) \cap I$. Then define $\delta = \min\{\delta_1, \delta_2\}$. For $x \in \mathsf{B}(\delta, x_0) \cap I$ we then have

$$\begin{aligned}\left|h(x)\Big(\frac{f(x)}{g(x)} - s_0\Big)\right| &= \left|h(x)\frac{f(x)}{g(x)} - h(x)s_0\right| \\ &\leq \left|h(x)\frac{f(x)}{g(x)} - s_0\right| + |(1 - h(x))s_0| \\ &< \frac{\epsilon}{2(1+|s_0|)} + \frac{\epsilon}{2(1+|s_0|)}|s_0| = \frac{\epsilon}{2}.\end{aligned}$$

Then, finally,

$$\left|\frac{f(x)}{g(x)} - s_0\right| < \frac{\epsilon}{2h(x)} < \epsilon,$$

for $x \in \mathsf{B}(\delta, x_0) \cap I$.

Now we consider the situation when $s_0 \in \{-\infty, \infty\}$. We shall take only the case of $s_0 = \infty$ since the other follows in a similar manner. We first take the case where $\lim_{x\uparrow x_0} f(x) = 0$ and $\lim_{x\uparrow x_0} g(x) = 0$. In this case, for $x \in I$, from the Cauchy Mean Value Theorem we can find $c_x \in (x, x_0)$ such that

$$\frac{f'(c_x)}{g'(c_x)} = \frac{f(x)}{g(x)}.$$

Now for $M \in \mathbb{R}_{>0}$ we choose $\delta \in \mathbb{R}_{>0}$ such that for $x \in \mathsf{B}(\delta, x_0) \cap I$ we have $\frac{f'(x)}{g'(x)} > M$. Then we immediately have

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)} > M$$

for $x \in \mathsf{B}(\delta, x_0) \cap I$ since $c_x \in \mathsf{B}(\delta, x_0)$, which gives the desired conclusion.

The final case we consider in this part of the proof is that where $s_0 = \infty$ and $\lim_{x \uparrow x_0} f(x) = \infty$ and $\lim_{x \uparrow x_0} g(x) = \infty$. For $M \in \mathbb{R}_{>0}$ choose $\delta_1 \in \mathbb{R}_{>0}$ such that $\frac{f'(x)}{g'(x)} > 2M$ provided that $x \in \mathsf{B}(\delta_1, x_0) \cap I$. Then, using Cauchy's Mean Value Theorem, for $x \in \mathsf{B}(\delta_1, x_0) \cap I$ there exists $c_x \in \mathsf{B}(\delta_1, x_0)$ such that

$$\frac{f'(c_x)}{g'(c_x)} = \frac{f(x) - f(x - \delta_1)}{g(x) - g(x - \delta_1)}.$$

Therefore,

$$\frac{f(x) - f(x - \delta_1)}{g(x) - g(x - \delta_1)} > 2M$$

for $x \in \mathsf{B}(\delta, x_0) \cap I$. As above, define

$$h(x) = \frac{1 - \frac{f(x-\delta_1)}{f(x)}}{1 - \frac{g(x-\delta_1)}{g(x)}}$$

and note that

$$\frac{f(x) - f(x - \delta_1)}{g(x) - g(x - \delta_1)} = h(x)\frac{f(x)}{g(x)}.$$

Therefore

$$h(x)\frac{f(x)}{g(x)} > 2M$$

for $x \in \mathsf{B}(\delta_1, x_0)$. Now take $\delta_2 \in \mathbb{R}_{>0}$ such that, if $x \in \mathsf{B}(\delta_2, x_0) \cap I$, then $h(x) \in [\frac{1}{2}, 2]$, this being possible since $\lim_{x \uparrow x_0} h(x) = 1$. It then follows that

$$\frac{f(x)}{g(x)} > \frac{2M}{h(x)} > M$$

for $x \in \mathsf{B}(\delta, x_0) \cap I$ where $\delta = \min\{\delta_1, \delta_2\}$.

(ii) This follows in the same manner as part (i).

(iii) This follows from parts (i) and (ii).

(iv) Let us define $\phi\colon (0, \infty) \to (0, \infty)$ by $\phi(x) = \frac{1}{x}$. Then define $\tilde{I} = \phi(I)$, noting that $\tilde{I}$ is an interval having 0 as an open left endpoint. Now define $\tilde{f}, \tilde{g}\colon \tilde{I} \to \mathbb{R}$ by $\tilde{f} = f \circ \phi$ and $\tilde{g} = g \circ \phi$. Using the Chain Rule (Theorem 3.2.13 below) we compute

$$\tilde{f}'(\tilde{x}) = f'(\phi(\tilde{x}))\phi'(\tilde{x}) = -\frac{f'(\frac{1}{\tilde{x}})}{\tilde{x}^2}$$

and similarly $\tilde{g}'(\tilde{x}) = -\frac{f'(\frac{1}{\tilde{x}})}{\tilde{x}^2}$. Therefore, for $\tilde{x} \in \tilde{I}$,

$$\frac{f'(\frac{1}{\tilde{x}})}{g'(\frac{1}{\tilde{x}})} = \frac{\tilde{f}'(\tilde{x})}{\tilde{g}'(\tilde{x})}.$$

and so, using part (ii) (it is easy to see that the hypotheses are verified),

$$\begin{aligned} &\lim_{\tilde{x}\downarrow 0}\frac{f'(\frac{1}{\tilde{x}})}{g'(\frac{1}{\tilde{x}})} = \lim_{\tilde{x}\downarrow 0}\frac{\tilde{f}'(\tilde{x})}{\tilde{g}'(\tilde{x})} \\ \implies \quad &\lim_{x\to\infty}\frac{f'(x)}{g'(x)} = \lim_{\tilde{x}\downarrow 0}\frac{\tilde{f}(\tilde{x})}{\tilde{g}(\tilde{x})} \\ \implies \quad &\lim_{x\to\infty}\frac{f'(x)}{g'(x)} = \lim_{x\to\infty}\frac{f(x)}{g(x)}, \end{aligned}$$

which is the desired conclusion.

(v) This follows in the same manner as part (iv). ■

3.2.22 Examples (Uses of L'Hôpital's Rule)

1. Let $I = \mathbb{R}$ and define $f, g\colon I \to \mathbb{R}$ by $f(x) = \sin x$ and $g(x) = x$. Note that f and g satisfy the hypotheses of Theorem 3.2.21 with $x_0 = 0$. Therefore we may compute
$$\lim_{x\to 0}\frac{f(x)}{g(x)} = \lim_{x\to 0}\frac{f'(x)}{g'(x)} = \frac{\cos 0}{1} = 1.$$
2. Let $I = [0, 1]$ and define $f, g\colon I \to \mathbb{R}$ by $f(x) = \sin x$ and $g(x) = x^2$. We can verify that f and g satisfy the hypotheses of L'Hôpital's Rule with $x_0 = 0$. Therefore we compute
$$\lim_{x\downarrow 0}\frac{f(x)}{g(x)} = \lim_{x\downarrow 0}\frac{f'(x)}{g'(x)} = \lim_{x\downarrow 0}\frac{\cos x}{2x} = \infty.$$
3. Let $I = \mathbb{R}_{>0}$ and define $f, g\colon I \to \mathbb{R}$ by $f(x) = e^x$ and $g(x) = -x$. Note that $\lim_{x\to\infty} f(x) = \infty$ and that $\lim_{x\to\infty} g(x) = -\infty$. Thus f and g do not quite satisfy the hypotheses of part (iv) of Theorem 3.2.21 since $\lim_{x\to\infty} g(x) \neq \infty$. However, the problem is a superficial one, as we now illustrate. Define $\tilde{g}(x) = -g(x) = x$. Then f and $\tilde{g}$ *do* satisfy the hypotheses of Theorem 3.2.21(iv). Therefore,
$$\lim_{x\to\infty}\frac{f(x)}{\tilde{g}(x)} = \lim_{x\to\infty}\frac{f'(x)}{\tilde{g}'(x)} = \lim_{x\to\infty}\frac{e^x}{1} = \infty,$$
and so
$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to\infty} -\frac{f(x)}{\tilde{g}(x)} = -\infty.$$
4. Consider the function $h\colon \mathbb{R} \to \mathbb{R}$ defined by $h(x) = \frac{x}{\sqrt{1+x^2}}$. We wish to determine $\lim_{x\to\infty} h(x)$, if this limit indeed exists. We will try to use L'Hôpital's Rule with $f(x) = x$ and $g(x) = \sqrt{1+x^2}$. First, one should check that f and g satisfy the hypotheses of the theorem taking $x_0 = 0$. One can check that f and g are differentiable on I and that $g'(x)$ is nonzero for $x \in I\backslash\{x_0\}$. Moreover, $\lim_{x\to 0} f(x) = 0$ and $\lim_{x\to 0} g(x) = 0$. Thus it only remains to check that $\lim_{x\to 0}\frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}$. To this end, one can easily compute that
$$\frac{f'(x)}{g'(x)} = \frac{g(x)}{f(x)},$$

which immediately implies that an application of L'Hôpital's Rule is destined to fail. However, the actual limit $\lim_{x\to\infty} h(x)$ does exist, however, and is readily computed, using the definition of limit, to be 1. Thus the converse of L'Hôpital's Rule does not hold. •

3.2.5 Monotonic functions and differentiability

In Section 3.1.5 we considered the notion of monotonicity, and its relationship with continuity. In this section we see how monotonicity is related to differentiability.

For functions that are differentiable, the matter of deciding on their monotonicity properties is straightforward.

3.2.23 Proposition (Monotonicity for differentiable functions) *For* $\mathrm{I} \subseteq \mathbb{R}$ *an interval and* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *a differentiable function, the following statements hold:*

(i) f *is constant if and only if* $\mathrm{f}'(\mathrm{x}) = 0$ *for all* $\mathrm{x} \in \mathrm{I}$*;*

(ii) f *is monotonically increasing if and only* $\mathrm{f}'(\mathrm{x}) \geq 0$ *for all* $\mathrm{x} \in \mathrm{I}$*;*

(iii) f *is strictly monotonically increasing if and only* $\mathrm{f}'(\mathrm{x}) > 0$ *for all* $\mathrm{x} \in \mathrm{I}$*;*

(iv) f *is monotonically decreasing if and only if* $\mathrm{f}'(\mathrm{x}) \leq 0$ *for all* $\mathrm{x} \in \mathrm{I}$*.*

(v) f *is strictly monotonically decreasing if and only if* $\mathrm{f}'(\mathrm{x}) < 0$ *for all* $\mathrm{x} \in \mathrm{I}$*.*

Proof In each case the "only if" assertions follow immediately from the definition of the derivative. To prove the "if" assertions, let $x_1, x_2 \in I$ with $x_1 < x_2$. By the Mean Value Theorem there exists $c \in [x_1, x_2]$ such that $f(x_1) - f(x_2) = f'(c)(x_1 - x_2)$. The result follows by considering the three cases of $f'(c) = 0$, $f'(c) \leq 0$, $f'(c) > 0$, $f'(c) \leq 0$, and $f'(c) < 0$, respectively. ■

The previous result gives the relationship between the derivative and monotonicity. Combining this with Theorem 3.1.30 which relates monotonicity with invertibility, we obtain the following characterisations of the derivative of the inverse function.

3.2.24 Theorem (Inverse Function Theorem for $\mathbb{R}$) *Let* $\mathrm{I} \subseteq \mathrm{J}$ *be an interval, let* $\mathrm{x}_0 \in \mathrm{I}$*, and let* $\mathrm{f}\colon \mathrm{I} \to \mathrm{J} = \mathrm{image}(\mathrm{f})$ *be a continuous, strictly monotonically increasing function that is differentiable at* x_0 *and for which* $\mathrm{f}'(\mathrm{x}_0) \neq 0$*. Then* $\mathrm{f}^{-1}\colon \mathrm{J} \to \mathrm{I}$ *is differentiable at* $\mathrm{f}(\mathrm{x}_0)$ *and the derivative is given by*

$$(\mathrm{f}^{-1})'(\mathrm{f}(\mathrm{x}_0)) = \frac{1}{\mathrm{f}'(\mathrm{x}_0)}.$$

Proof From Theorem 3.1.30 we know that f is invertible. Let $y_0 = f(x_0)$, let $y_1 \in J$, and define $x_1 \in I$ by $f(x_1) = y_1$. Then, if $x_1 \neq x_0$,

$$\frac{f^{-1}(y_1) - f^{-1}(y_0)}{y_1 - y_0} = \frac{x_1 - x_0}{f(x_1) - f(x_0)}.$$

Therefore,

$$(f^{-1})'(y_0) = \lim_{y_1 \to_J y_0} \frac{f^{-1}(y_1) - f^{-1}(y_0)}{y_1 - y_0} = \lim_{x_1 \to_I x_0} \frac{x_1 - x_0}{f(x_1) - f(x_0)} = \frac{1}{f'(x_0)},$$

as desired. ■

3.2.25 Corollary (Alternate version of Inverse Function Theorem) *Let* $I \subseteq \mathbb{R}$ *be an interval, let* $x_0 \in I$*, and let* $f\colon I \to \mathbb{R}$ *be a function of class* C^1 *such that* $f'(x_0) \neq 0$*. Then there exists a neighbourhood* U *of* x_0 *in* I *and a neighbourhood* V *of* $f(x_0)$ *such that* $f|U$ *is invertible, and such that* $(f|U)^{-1}$ *is differentiable, and the derivative is given by*

$$((f|U)^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

for each $y \in V$.

Proof Since f' is continuous and is nonzero at x_0, there exists a neighbourhood U of x_0 such that $f'(x)$ has the same sign as $f'(x_0)$ for all $x \in U$. Thus, by Proposition 3.2.23, $f|U$ is either strictly monotonically increasing (if $f'(x_0) > 0$) or strictly monotonically decreasing (if $f'(x_0) < 0$). The result now follows from Theorem 3.2.24. ■

For general monotonic functions, Proposition 3.2.23 turns out to be "almost" enough to characterise them. To understand this, we recall from Section 2.5.6 the notion of a subset of $\mathbb{R}$ of measure zero. With this recollection having been made, we have the following characterisation of general monotonic functions.

3.2.26 Theorem (Characterisation of monotonic functions II) *If* $I \subseteq \mathbb{R}$ *is an interval and if* $f\colon I \to \mathbb{R}$ *is either monotonically increasing (resp. monotonically decreasing), then* f *is differentiable almost everywhere, and* $f'(x) \geq 0$ *(resp.* $f'(x) \leq 0$*) at all points* $x \in I$ *where* f *is differentiable.*

Proof We first prove a technical lemma.

1 Lemma *If* $g\colon [a,b] \to \mathbb{R}$ *has the property that, for each* $x \in [a,b]$*, the limits* $g(x+)$ *and* $g(x-)$ *exist whenever they are defined as limits in* $[a,b]$*. If we define*

$$S = \{x \in [a,b] \mid \textit{there exists } x' > x \textit{ such that } g(x') > \max\{g(x-), g(x), g(x+)\}\},$$

then S *is a disjoint union of a countable collection* $\{I_\alpha \mid \alpha \in A\}$ *of intervals that are open as subsets of* $[a,b]$ *(cf. the beginning of Section 3.1.1).*

Proof Let $x \in S$. We have three cases.

1. There exists $x' > x$ such that $g(x') > g(x-)$, and $g(x-) \geq g(x)$ and $g(x-) \geq g(x+)$: Define $g_{x,-}, g_{x,+}\colon [a,b] \to \mathbb{R}$ by

$$g_{x,-}(y) = \begin{cases} g(y), & y \neq 1, \\ g(x-), & y = x, \end{cases} \qquad g_{x,+}(y) = \begin{cases} g(y), & y \neq 1, \\ g(x+), & y = x. \end{cases}$$

Since the limit $g(x-)$ exists, $g_{x,-}|[a,x]$ is continuous at x by Theorem 3.1.3. Since $g(x') > g_{x,-}(x)$, there exists $\epsilon_1 \in \mathbb{R}_{>0}$ such that $g(x') > g_{x,-}(y) = g(y)$ for all $y \in (x - \epsilon_1, x)$. Now note that $g(x') > g(x-) \geq g_{x,+}(x)$. Arguing similarly to what we have done, there exists $\epsilon_2 \in \mathbb{R}_{>0}$ such that $g(x') > g_{x,+}(y) = g(y)$ for all $y \in (x, x+\epsilon_2)$. Let $\epsilon = \min\{\epsilon_1, \epsilon_2\}$. Since $g(x') > g(x-) \geq g(x)$, it follows that $g(x') > g(y)$ for all $y \in (x - \epsilon, x + \epsilon)$, so we can conclude that S is open.

2. There exists $x' > x$ such that $g(x') > g(x)$, and $g(x) \geq g(x-)$ and $g(x) \geq g(x+)$: Define $g_{x,-}$ and $g_{x,+}$ as above. Then, since $g(x') > g(x) \geq g(x-)$ and $g(x') > g(x) \geq g(x+)$, we can argue as in the previous case that there exists $\epsilon \in \mathbb{R}_{>0}$ such that $g(x') > g(y)$ for all $y \in (x - \epsilon, x + \epsilon)$. Thus S is open.

3. There exists $x' > x$ such that $g(x') > g(x+)$, and $g(x+) \geq g(x)$ and $g(x+) \geq g(x-)$: Here we can argue in a manner entirely similar to the first case that S is open.

The preceding arguments show that S is open, and so by Proposition 2.5.6 it is a countable union of open intervals. ▼

Now define

$$\Lambda_l(x) = \limsup_{h \downarrow 0} \frac{f(x-h) - f(x)}{-h} \qquad \lambda_l(x) = \liminf_{h \downarrow 0} \frac{f(x-h) - f(x)}{-h}$$
$$\Lambda_r(x) = \limsup_{h \downarrow 0} \frac{f(x+h) - f(x)}{h} \qquad \lambda_r(x) = \liminf_{h \downarrow 0} \frac{f(x+h) - f(x)}{h}.$$

If f is differentiable at x then these four numbers will be finite and equal. We shall show that

1. $\Lambda_r(x) < \infty$ and
2. $\Lambda_r(x) \leq \lambda_l(x)$

for almost every $x \in [a, b]$. Since the relations

$$\lambda_l \leq \Lambda_l \leq \lambda_r \leq \Lambda_r$$

hold due to monotonicity of f, the differentiability of f for almost all x will then follow.

For 1, if $M \in \mathbb{R}_{>0}$ denote

$$S_M = \{x \in [a, b] \mid \Lambda_r(x) > M\}.$$

Thus, for $x_0 \in S_M$, there exists $x > x_0$ such that

$$\frac{f(x) - f(x_0)}{x - x_0} > M.$$

Defining $g_M(x) = f(x) - Mx$ this asserts that $g_M(x) > g_M(x_0)$. The function g_M satisfies the hypotheses of Lemma 1 by part (i). This means that S_M is contained in a finite or countable disjoint union of intervals $\{I_\alpha \mid \alpha \in A\}$, open in $[a, b]$, for which

$$g_M(a_\alpha) \leq \max\{g_M(b_\alpha-), g_M(b_\alpha), g_M(b_\alpha+)\}, \qquad \alpha \in A,$$

where a_α and b_α are the left and right endpoints, respectively, for I_α, $\alpha \in A$. In particular, $g_M(a_\alpha) \leq g_M(b_\alpha)$. A trivial manipulation then gives

$$M(b_\alpha - a_\alpha) \leq f(b_\alpha) - f(a_\alpha), \qquad \alpha \in A.$$

We have

$$M \sum_{\alpha \in A} |b_\alpha - a_\alpha| \leq \sum_{\alpha \in A} |f(b_\alpha) - f(a_\alpha)| \leq f(b) - f(a)$$

since f is monotonically increasing. Since f is bounded, this shows that as $M \to \infty$ the length of the open intervals $\{(a_\alpha, b_\alpha) \mid \alpha \in A\}$ covering S_M must go to zero. This shows that the set of points where 1 holds has zero measure.

Now we turn to 2. Let $0 < m < M$, define $g_m(x) = -f(x) + mx$ and $g_M(x) = f(x) - Mx$. Also define

$$S_m = \{x \in [a, b] \mid \lambda_l(x) < m\}.$$

For $x_0 \in S_m$ there exists $x < x_0$ such that

$$\frac{f(x) - f(x_0)}{x - x_0} < m,$$

which is equivalent to $g_m(x) > g_m(x_0)$. Therefore, by Lemma 1, note that S_m is contained in a finite or countable disjoint union of intervals $\{I_\alpha \mid \alpha \in A\}$, open in $[a, b]$. Denote by a_α and b_α the left and right endpoints, respectively, for I_α for $\alpha \in A$. For $\alpha \in A$ denote

$$S_{\alpha,M} = \{x \in [a_\alpha, b_\alpha] \mid \Lambda_r(x) > M\},$$

and arguing as we did in the proof that 1 holds almost everywhere, denote by $\{I_{\alpha,\beta} \mid \beta \in B_\alpha\}$ the countable collection of subintervals, open in $[a, b]$, of (a_α, b_α) that contain $S_{\alpha,M}$. Denote by $a_{\alpha,\beta}$ and $b_{\alpha,\beta}$ the left and right endpoints, respectively, of $I_{\alpha,\beta}$ for $\alpha \in A$ and $\beta \in B_\alpha$. Note that the relations

$$\begin{aligned}
&g_m(a_\alpha) \le \max\{g_m(b_\alpha-), g_m(b_\alpha), g_m(b_\alpha+)\}, \qquad \alpha \in A,\\
&g_M(a_{\alpha,\beta}) \le \max\{g_M(b_{\alpha,\beta}-), g_M(b_{\alpha,\beta}), g_M(b_{\alpha,\beta}+)\}, \qquad \alpha \in A,\ \beta \in B_\alpha
\end{aligned}$$

hold. We then may easily compute

$$\begin{aligned}
&f(b_\alpha) - f(a_\alpha) \le m(b_\alpha - a_\alpha), \qquad \alpha \in A,\\
&f(b_{\alpha,\beta}) - f(a_{\alpha,\beta}) \ge M(b_{\alpha,\beta} - b_{\alpha,\beta}), \qquad \alpha \in A,\ \beta \in A_\alpha.
\end{aligned}$$

Therefore, for each $\alpha \in A$,

$$M \sum_{\beta \in A_\alpha} |b_{\alpha,\beta} - a_{\alpha,\beta}| \le \sum_{\beta \in A_\alpha} |f(b_{\alpha,\beta} - a_{\alpha,\beta})| \le f(b_\alpha) - f(a_\alpha) \le m(b_\alpha - a_\alpha).$$

This then gives

$$M \sum_{\alpha \in A} \sum_{\beta \in A_\alpha} |b_{\alpha,\beta} - a_{\alpha,\beta}| \le m \sum_{\alpha \in A} |b_\alpha - a_\alpha|,$$

or $\Sigma_2 \le \frac{m}{M}\Sigma_1$, where

$$\Sigma_1 = \sum_{\alpha \in A} \sum_{\beta_\alpha \in K_\alpha} |b_{\alpha,\beta} - a_{\alpha,\beta}|, \quad \Sigma_2 = \sum_{\alpha \in A} |b_\alpha - a_\alpha|.$$

Now, this process can be repeated, defining

$$S_{\alpha,\beta,m} = \{x \in [a_{\alpha,\beta}, b_{\alpha,\beta}] \mid \lambda_l(x) < m\},$$

and so on. We then generate a sequence of finite or countable disjoint intervals of total length Σ_α and satisfying

$$\Sigma_{2\alpha} \le \frac{m}{M}\Sigma_{2\alpha-1} \le \Big(\frac{m}{M}\Big)^\alpha \Sigma_1, \qquad \alpha \in A.$$

It therefore follows that $\lim_{\alpha\to\infty} \Sigma_\alpha = 0$. Thus the set of points

$$S_{M,m} = \{x \in [a, b] \mid m < \lambda_l(x) \text{ and } \Lambda_r(x) > M\}$$

is contained in a set of zero measure provided that $m < M$. Now note that

$$\{x \in [a,b] \mid \lambda_l(x) \geq \Lambda_r(x)\} \subseteq \bigcup \{S_{M,m} \mid m,M \in \mathbb{Q},\ m < M\}.$$

The union on the left is a countable union of sets of zero measure, and so has zero measure itself (by Exercise 2.5.9). This shows that f is differentiable on a set whose complement has zero measure.

To show that $f'(x) \geq 0$ for all points x at which f is differentiable, suppose the converse. Thus suppose that there exists $x \in [a,b]$ such that $f'(x) < 0$. This means that for ϵ sufficiently small and positive,

$$\frac{f(x+\epsilon)-f(x)}{\epsilon} < 0 \quad \Longrightarrow \quad f(x+\epsilon)-f(x) < 0,$$

which contradicts the fact that f is monotonically increasing. This completes the proof of the theorem. ■

Let us give two examples of functions that illustrate the surprisingly strange behaviour that can arise from monotonic functions. These functions are admittedly degenerate, and not something one is likely to encounter in applications. However, they do show that one cannot strengthen the conclusions of Theorem 3.2.26.

Our first example is one of the standard "peculiar" monotonic functions, and its construction relies on the middle-thirds Cantor set constructed in Example 2.5.39.

3.2.27 Example (A continuous increasing function with an almost everywhere zero derivative) Let C_k, $k \in \mathbb{Z}_{>0}$, be the sets, comprised of collections of disjoint closed intervals, used in the construction of the middle-thirds Cantor set of Example 2.5.39. Note that, for $x \in [0,1]$, the set $[0,x] \cap C_k$ consists of a finite number of intervals. Let $g_k\colon [0,1] \to [0,1]$ be defined by asking that $g_{C,k}(x)$ be the sum of the lengths of the intervals comprising $[0,x] \cap C_k$. Then define $f_{C,k}\colon [0,1] \to [0,1]$ by $f_{C,k}(x) = \left(\frac{3}{2}\right)^k g_{C,k}(x)$. Thus $f_{C,k}$ is a function that is constant on the complement to the closed intervals comprising C_k, and is linear on those same closed intervals, with a slope determined in such a way that the function is continuous. We then define $f_C\colon [0,1] \to [0,1]$ by $f_C(x) = \lim_{k\to\infty} f_{C,k}(x)$. In Figure 3.9 we depict f_C. The reader new to this function should take the requisite moment or two to understand our definition of f_C, perhaps by sketching a couple of the functions $f_{C,k}$, $k \in \mathbb{Z}_{>0}$.

Let us record some properties of the function f_C, which is called the ***Cantor function*** or the ***Devil's staircase***.

1 Lemma f_C *is continuous.*

Proof We prove this by showing that the sequence of functions $(f_{C,k})_{k\in\mathbb{Z}_{>0}}$ converges uniformly, and then using Theorem 3.5.8 to conclude that the limit function is continuous. Note that the functions $f_{C,k}$ and $f_{C,k+1}$ differ only on the closed intervals comprising C_k. Moreover, if $J_{k,j}$, $k \in \mathbb{Z}_{\geq 0}$, $j \in \{1,\dots,2^k-1\}$, denotes the set of open intervals forming $[0,1] \setminus C_k$, numbered from left to right, then the value of $f_{C,k}$ on $J_{k,j}$ is $j2^{-k}$. Therefore,

$$\sup\{|f_{C,k+1}(x) - f_{C,k}(x)| \mid x \in [0,1]\} < 2^{-k}, \qquad k \in \mathbb{Z}_{\geq 0}.$$

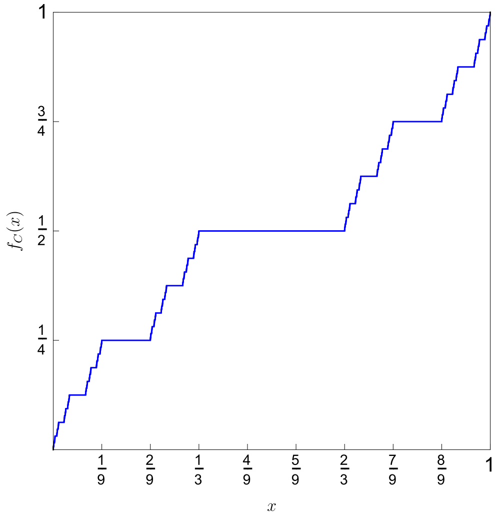


Figure 3.9 A depiction of the Cantor function

This implies that $(f_{C,k})_{k\in\mathbb{Z}_{>0}}$ is uniformly convergent as in Definition 3.5.4. Thus Theorem 3.5.8 gives continuity of f_C, as desired. ▼

2 Lemma f_C *is differentiable at all points in* $[0,1]\setminus C$*, and its derivative, where it exists, is zero.*

Proof Since C is constructed as an intersection of the closed sets C_k, and since such intersections are themselves closed by Exercise 2.5.1, it follows that $[0,1]\setminus C$ is open. Thus if $x\in[0,1]\setminus C$, there exists $\epsilon\in\mathbb{R}_{>0}$ such that $\mathsf{B}(\epsilon,x)\subseteq[0,1]\setminus C$. Since $\mathsf{B}(\epsilon,x)$ contains no endpoints for intervals from the sets C_k, $k\in\mathbb{Z}_{>0}$, it follows that $f_{C,k}|\mathsf{B}(\epsilon,x)$ is constant for sufficiently large k. Therefore $f_C|\mathsf{B}(\epsilon,x)$ is constant, and it then follows that f_C is differentiable at x, and that $f_C'(x)=0$. ▼

In Example 2.5.39 we showed that C has measure zero. Thus we have a continuous, monotonically increasing function from $[0,1]$ to $[0,1]$ whose derivative is almost everywhere zero. It is perhaps not *a priori* obvious that such a function can exist, since one's first thought might be that zero derivative implies a constant function. The reasons for the failure of this rule of thumb in this example will not become perfectly clear until we examine the notion of absolute continuity in Section 5.9.6. •

The second example of a "peculiar" monotonic function is not quite as standard in the literature, but is nonetheless interesting since it exhibits somewhat different oddities than the Cantor function.

3.2.28 Example (A strictly increasing function, discontinuous on the rationals, with an almost everywhere zero derivative) We define a strictly monotonically increasing function $f_{\mathbb{Q}}\colon\mathbb{R}\to\mathbb{R}$ as follows. Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be an enumeration of the rational numbers and for $x \in \mathbb{R}$ define

$$I(x) = \{j \in \mathbb{Z}_{>0} \mid q_j < x\}.$$

Now define

$$f_{\mathbb{Q}}(x) = \sum_{j \in I(x)} \frac{1}{2^j}.$$

Let us record the properties of $f_{\mathbb{Q}}$ in a series of lemmata.

1 Lemma $\lim_{x \to -\infty} f_{\mathbb{Q}}(x) = 0$ *and* $\lim_{x \to \infty} f_{\mathbb{Q}}(x) = 1$.

Proof Recall from Example 2.4.2–1 that $\sum_{j=1}^{\infty} \frac{1}{2^j} = 1$. Let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $\sum_{j=N+1}^{\infty} \frac{1}{2^j} < \epsilon$. Now choose $M \in \mathbb{R}_{>0}$ such that $\{q_1, \dots, q_N\} \subseteq [-M, M]$. Then, for $x < M$ we have

$$f_{\mathbb{Q}}(x) = \sum_{j \in I(x)} \frac{1}{2^j} = \sum_{j=1}^{\infty} \frac{1}{2^j} - \sum_{j \in \mathbb{Z}_{>0} \setminus I(x)} \frac{1}{2^j} \leq \sum_{j=1}^{\infty} \frac{1}{2^j} - \sum_{j=1}^{N} \frac{1}{2^j} < \epsilon.$$

Also, for $x > M$ we have

$$f_{\mathbb{Q}}(x) = \sum_{j \in I(x)} \frac{1}{2^j} \geq \sum_{j=1}^{N} \frac{1}{2^j} > 1 - \epsilon.$$

Thus $\lim_{x \to -\infty} f_{\mathbb{Q}}(x) = 0$ and $\lim_{x \to \infty} f_{\mathbb{Q}}(x) = 1$. ▼

2 Lemma $f_{\mathbb{Q}}$ *is strictly monotonically increasing.*

Proof Let $x, y \in \mathbb{R}$ with $x < y$. Then, by Corollary 2.2.16, there exists $q \in \mathbb{Q}$ such that $x < q < y$. Let $j_0 \in \mathbb{Z}_{>0}$ have the property that $q = q_{j_0}$. Then

$$f_{\mathbb{Q}}(y) = \sum_{j \in I(y)} \frac{1}{2^j} \geq \sum_{j \in I(x)} \frac{1}{2^j} + \frac{1}{2^{j_0}} > f_{\mathbb{Q}}(x),$$

as desired. ▼

3 Lemma $f_{\mathbb{Q}}$ *is discontinuous at each point in* $\mathbb{Q}$.

Proof Let $q \in \mathbb{Q}$ and let $x > q$. Let $j_0 \in \mathbb{Z}_{>0}$ satisfy $q = q_{j_0}$. Then

$$f_{\mathbb{Q}}(x) = \sum_{j \in I(x)} \frac{1}{2^j} \geq \frac{1}{2^{j_0}} + \sum_{j \in I(q)} \frac{1}{2^j} = \frac{1}{2^{j_0}} + \sum_{j \in I(q)} \frac{1}{2^j}.$$

Therefore, $\lim_{x \downarrow q} f_{\mathbb{Q}}(x) \geq \frac{1}{2^{j_0}} + f_{\mathbb{Q}}(q)$, implying that $f_{\mathbb{Q}}$ is discontinuous at q by Theorem 3.1.3. ▼

4 Lemma $f_{\mathbb{Q}}$ *is continuous at each point in* $\mathbb{R} \setminus \mathbb{Q}$*.*

Proof Let $x \in \mathbb{R} \setminus \mathbb{Q}$ and let $\epsilon \in \mathbb{R}_{>0}$. Take $N \in \mathbb{Z}_{>0}$ such that $\sum_{j=N+1}^{\infty} \frac{1}{2^j} < \epsilon$ and define $\delta \in \mathbb{R}_{>0}$ such that $\mathsf{B}(\delta, x) \cap \{q_1, \dots, q_N\} = \emptyset$ (why is this possible?). Now let

$$I(\delta, x) = \{j \in \mathbb{Z}_{>0} \mid q_j \in \mathsf{B}(\delta, x)\}$$

and note that, for $y \in \mathsf{B}(\delta, x)$ with $x < y$, we have

$$\begin{aligned} f_{\mathbb{Q}}(y) - f_{\mathbb{Q}}(x) = \sum_{j \in I(y)} \frac{1}{2^j} - \sum_{j \in I(x)} \frac{1}{2^j} \le \sum_{j \in I(\delta,x)} \frac{1}{2^j} = \sum_{j=1}^{\infty} \frac{1}{2^j} - \sum_{\mathbb{Z}_{>0} \setminus I(\delta,x)} \frac{1}{2^j} \\ \le \sum_{j=1}^{\infty} \frac{1}{2^j} - \sum_{j=1}^{N} \frac{1}{2^j} = \sum_{j=N+1}^{\infty} \frac{1}{2^j} < \epsilon. \end{aligned}$$

A similar argument holds for $y < x$ giving $f_{\mathbb{Q}}(x) - f_{\mathbb{Q}}(y) < \epsilon$ in this case. Thus $|f_{\mathbb{Q}}(y) - f_{\mathbb{Q}}(x)| < \epsilon$ for $|y - x| < \delta$, thus showing continuity of f at x. ▼

5 Lemma *The set* $\{x \in \mathbb{R} \mid f'_{\mathbb{Q}}(x) \neq 0\}$ *has measure zero.*

Proof The proof relies on some concepts from Section 3.5. For $k \in \mathbb{Z}_{>0}$ define $f_{\mathbb{Q},k} \colon \mathbb{R} \to \mathbb{R}$ by

$$f_{\mathbb{Q},k}(x) = \sum_{j \in I(x) \cap \{1,\dots,k\}} \frac{1}{2^j}.$$

Note that $(f_{\mathbb{Q},k})_{k \in \mathbb{Z}_{>0}}$ is a sequence of monotonically increasing functions with the following properties:

1. $\lim_{k \to \infty} f_{\mathbb{Q},k}(x) = f_{\mathbb{Q}}(x)$ for each $x \in \mathbb{R}$;
2. the set $\{x \in \mathbb{R} \mid f'_{\mathbb{Q},k}(x) \neq 0\}$ is finite for each $k \in \mathbb{Q}$.

The result now follows from Theorem 3.5.25. ▼

Thus we have an example of a strictly monotonically increasing function whose derivative is zero almost everywhere. Note that this function also has the feature that in any neighbourhood of a point where it is differentiable, there lie points where it is not differentiable. This is an altogether peculiar function. •

3.2.6 Convex functions and differentiability

Let us now return to our consideration of convex functions introduced in Section 3.1.6. Here we discuss the differentiability properties of convex functions. The following notation for a function $f \colon I \to \mathbb{R}$ will be convenient:

$$f'(x+) = \lim_{\epsilon \downarrow 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}, \quad f'(x-) = \lim_{\epsilon \downarrow 0} \frac{f(x) - f(x-\epsilon)}{\epsilon},$$

provided that these limits exist.

With this notation, convex functions have the following properties.

3.2.29 Proposition (Properties of convex functions II) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$ *and for a convex function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$*, the following statements hold:*

(i) if I *is open then the limits* $\mathrm{f}'(\mathrm{x}+)$ *and* $\mathrm{f}'(\mathrm{x}-)$ *exist and* $\mathrm{f}'(\mathrm{x}-) \le \mathrm{f}'(\mathrm{x}+)$ *for each* $\mathrm{x} \in \mathrm{I}$*;*

(ii) if I *is open then the functions*

$$\mathrm{I} \ni \mathrm{x} \mapsto \mathrm{f}'(\mathrm{x}+), \quad \mathrm{I} \ni \mathrm{x} \mapsto \mathrm{f}'(\mathrm{x}-)$$

are monotonically increasing, and strictly monotonically increasing if f *is strictly convex;*

(iii) if I *is open and if* $\mathrm{x}_1, \mathrm{x}_2 \in \mathrm{I}$ *satisfy* $\mathrm{x}_1 < \mathrm{x}_2$*, then* $\mathrm{f}'(\mathrm{x}_1+) \le \mathrm{f}'(\mathrm{x}_2-)$*;*

(iv) f *is differentiable except at a countable number of points in* I*.*

Proof (i) Since I is open there exists $\epsilon_0 \in \mathbb{R}_{>0}$ such that $[x, x+\epsilon_0) \subseteq I$. Let $(\epsilon_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $(0, \epsilon_0)$ converging to 0 and such that $\epsilon_{j+1} < \epsilon_j$ for every $j \in \mathbb{Z}_{>0}$. Then the sequence $(s_f(x, x+\epsilon_j))_{j\in\mathbb{Z}_{>0}}$ is monotonically decreasing. This means that, by Lemma 3.1.33,

$$\frac{f(x+\epsilon_{j+1}) - f(x)}{\epsilon_{j+1}} \le \frac{f(x+\epsilon_j) - f(x)}{\epsilon_j}$$

for each $j \in \mathbb{Z}_{>0}$. Moreover, if $x' \in I$ satisfies $x' < x$ then we have $s_f(x', x) \le s_f(x, x+\epsilon_j)$ for each $j \in \mathbb{Z}_{>0}$. Thus the sequence $(\epsilon_j^{-1}(f(x+\epsilon_j) - f(x)))_{j\in\mathbb{Z}_{>0}}$ is decreasing and bounded from below. Thus it must converge, cf. Theorem 2.3.8.

The proof for the existence of the other asserted limit follows that above, *mutatis mutandis*.

To show that $f'(x-) \le f'(x+)$, note that, for all ϵ sufficiently small,

$$\frac{f(x) - f(x-\epsilon)}{\epsilon} = s_f(x-\epsilon, x) \le s_f(x, x+\epsilon) = \frac{f(x+\epsilon) - f(x)}{\epsilon}.$$

Taking limits as $\epsilon \downarrow 0$ gives the desired inequality.

(ii) For $x_1, x_2 \in I$ with $x_1 < x_2$ we have

$$f'(x_1+) = \lim_{\epsilon\downarrow 0} s_f(x_1, x_1+\epsilon) \le \lim_{\epsilon\downarrow 0} s_f(x_2, x_2+\epsilon) = f'(x_2+),$$

using Lemma 3.1.33. A similar computation, *mutatis mutandis*, shows that the other function in this part of the result is also monotonically increasing. Moreover, if f is strictly convex that the inequalities above can be replaced with strict inequalities by (3.2). From this we conclude that $x \mapsto f'(x_+)$ and $x \mapsto f'(x_-)$ are strictly monotonically increasing.

(iii) For $\epsilon \in \mathbb{R}_{>0}$ sufficiently small we have

$$x_1 + \epsilon < x_2 - \epsilon.$$

For all such sufficiently small ϵ we have

$$\frac{f(x_1+\epsilon) - f(x_1)}{\epsilon} = s_f(x_1, x_1+\epsilon) \le s_f(x_2-\epsilon, x_2) = \frac{f(x_2) - f(x_2-\epsilon)}{\epsilon}$$

by Lemma 3.1.33. Taking limits as $\epsilon \downarrow 0$ gives this part of the result.

(iv) Let A_f be the set of points in I where f is not differentiable. Note that

$$\frac{f(x)-f(x-\epsilon)}{\epsilon} = s_f(x-\epsilon, x) \leq s_f(x, x+\epsilon) = \frac{f(x+\epsilon)-f(x)}{\epsilon}$$

by Lemma 3.1.33. Therefore, if $x \in A_f$, then $f'(x-) < f'(x+)$. We define a map $\phi\colon A_f \to \mathbb{Q}$ as follows. If $x \in A_f$ we use the Axiom of Choice and Corollary 2.2.16 to select $\phi(x) \in \mathbb{Q}$ such that $f'(x-) < \phi(x) < f'(x+)$. We claim that ϕ is injective. Indeed, if $x, y \in A_f$ are distinct (say $x < y$) then, using parts (ii) and (iii),

$$f'(x-) < \phi(x) < f'(x+) < f'(y-) < \phi(y) < f'(y+).$$

Thus $\phi(x) < \phi(y)$ and so ϕ is injective as desired. Thus A_f must be countable. ■

For functions that are sufficiently differentiable, it is possible to conclude convexity from properties of the derivative.

3.2.30 Proposition (Convexity and derivatives) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$ *and for a function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *the following statements hold:*

(i) for each $\mathrm{x}_1, \mathrm{x}_2 \in \mathrm{I}$ *with* $\mathrm{x}_1 \neq \mathrm{x}_2$ *we have*

$$\mathrm{f}(\mathrm{x}_2) \geq \mathrm{f}(\mathrm{x}_1) + \mathrm{f}'(\mathrm{x}_1+)(\mathrm{x}_2 - \mathrm{x}_1), \quad \mathrm{f}(\mathrm{x}_2) \geq \mathrm{f}(\mathrm{x}_1) + \mathrm{f}'(\mathrm{x}_1-)(\mathrm{x}_2 - \mathrm{x}_1);$$

(ii) if f *is differentiable, then* f *is convex if and only if* f′ *is monotonically increasing;*

(iii) if f *is differentiable, then* f *is strictly convex if and only if* f′ *is strictly monotonically increasing;*

(iv) if f *is twice continuously differentiable, then it is convex if and only if* $\mathrm{f}''(\mathrm{x}) \geq 0$ *for every* $\mathrm{x} \in \mathrm{I}$*;*

(v) if f *is twice continuously differentiable, then it is strictly convex if and only if* $\mathrm{f}''(\mathrm{x}) > 0$ *for every* $\mathrm{x} \in \mathrm{I}$.

Proof (i) Suppose that $x_1 < x_2$. Then, for $\epsilon \in \mathbb{R}_{>0}$ sufficiently small,

$$\frac{f(x_1+\epsilon)-f(x_1)}{\epsilon} \leq \frac{f(x_2)-f(x_1)}{x_2 - x_1}$$

by Lemma 3.1.33. Thus, taking limits as $\epsilon \downarrow 0$,

$$f'(x_1+) \leq \frac{f(x_2)-f(x_1)}{x_2-x_1},$$

and rearranging gives

$$f(x_2) \geq f(x_1) + f'(x_1+)(x_2 - x_1).$$

Since we also have $f'(x_1-) \leq f'(x_1+)$ by Proposition 3.2.29(i), we have both of the desired inequalities in this case.

Now suppose that $x_2 < x_1$. Again, for $\epsilon \in \mathbb{R}_{>0}$ sufficiently small, we have

$$\frac{f(x_1+\epsilon)-f(x_1)}{\epsilon} \geq \frac{f(x_1)-f(x_2)}{x_1-x_2},$$

and taking the limit as $\epsilon \downarrow 0$ gives

$$f'(x_1+) \geq \frac{f(x_1) - f(x_2)}{x_1 - x_2}.$$

Rearranging gives

$$f(x_2) \geq f(x_1) + f'(x_1+)(x_2 - x_1)$$

and since $f'(x_1-) \leq f'(x_1+)$ the desired inequalities follow in this case.

(ii) From Proposition 3.2.29(ii) we deduce that if f is convex and differentiable then f' is monotonically increasing. Conversely, suppose that f is differentiable and that f' is monotonically increasing. Let $x_1, x_2 \in I$ satisfy $x_1 < x_2$ and let $s \in (0, 1)$. By the Mean Value Theorem there exists $c_1, c_2 \in I$ satisfying

$$x_1 < c_1 < (1 - s)x_1 + sx_2 < d_1 < x_2$$

such that

$$\frac{f((1-s)x_1 + sx_2) - f(x_1)}{(1-s)x_1 + sx_2 - x_1} = f'(c_1) \leq f'(c_2) = \frac{f(x_2) - f((1-s)x_1 + sx_2)}{x_2 - ((1-s)x_1 + sx_2)}. \tag{3.9}$$

Rearranging, we get

$$\frac{f((1-s)x_1 + sx_2) - f(x_1)}{s(x_2 - x_1)} \leq \frac{f(x_2) - f((1-s)x_1 + sx_2)}{(1-s)(x_2 - x_1)},$$

and further rearranging gives

$$f((1-s)x_1 + sx_2) \leq (1-s)f(x_1) + sf(x_2),$$

and so f is convex.

(iii) If f is strictly convex, then from Proposition 3.2.29 we conclude that f' is strictly monotonically increasing. Next suppose that f' is strictly monotonically decreasing and let $x_1, x_2 \in I$ satisfy $x_1 < x_2$ and let $s \in (0, 1)$. The proof that f is strictly convex follows as in the preceding part of the proof, noting that, in (3.9), we have $f'(c_1) < f'(c_2)$. Carrying this strict inequality through the remaining computations shows that

$$f((1-s)x_1 + sx_2) \leq (1-s)f(x_1) + sf(x_2),$$

giving strict convexity of f.

(iv) If f'' is nonnegative, then f' is monotonically increasing by Proposition 3.2.23. The result now follows from part (ii).

(iv) If f'' is positive, then f' is strictly monotonically increasing by Proposition 3.2.23. The result now follows from part (iii). ■

Let us consider a few examples illustrating how convexity and differentiability are related.

3.2.31 Examples (Convex functions and differentiability)

1. The convex function $n_{x_0}\colon \mathbb{R} \to \mathbb{R}$ defined by $n_{x_0}(x) = |x - x_0|$ is differentiable everywhere except for $x = x_0$. But at $x = x_0$ the derivatives from the left and right exist. Moreover, $f'(x) = -1$ for $x < x_0$ and $f'(x) = 1$ for $x > x_0$. Thus we see that the derivative is monotonically increasing, although it is not defined everywhere.
2. As we showed in Proposition 3.2.29(iv), a convex function is differentiable except at a countable set of points. Let us show that this conclusion cannot be improved. Let $C \subseteq \mathbb{R}$ be a countable set. We shall construct a convex function $f\colon \mathbb{R} \to \mathbb{R}$ whose derivative exists on $\mathbb{R} \setminus C$ and does not exist on C. In case C is finite, we write $C = \{x_1, \dots, x_k\}$. Then one verifies that the function f defined by

$$f(x) = \sum_{j=1}^{k} |x - x_j|$$

is verified to be convex, being a finite sum of convex functions (see Proposition 3.1.39). It is clear that f is differentiable at points in $\mathbb{R} \setminus C$ and is not differentiable at points in C. Now suppose that C is not finite. Let us write $C = \{x_j\}_{j \in \mathbb{Z}_{>0}}$, i.e., enumerate the points in C. Let us define $c_j = (2^j \max\{1, |x_j|\})^{-1}$, $j \in \mathbb{Z}_{>0}$, and define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \sum_{j=1}^{\infty} c_j |x - x_j|.$$

We shall prove that this function is well-defined, convex, differentiable at points in $\mathbb{R} \setminus C$, and not differentiable at points in C. In proving this, we shall make reference to some results we have not yet proved.

First let us show that f is well-defined.

1 Lemma *For every compact subset* $K \subseteq \mathbb{R}$*, the series*

$$\sum_{j=1}^{\infty} \mathrm{c_j|x - x_j|}$$

converges uniformly on K *(see Section 3.5.2 for uniform convergence).*

Proof Let $K \subseteq \mathbb{R}$ and let $R \in \mathbb{R}_{>0}$ be large enough that $K \subseteq [-R, R]$. Then, for $x \in K$ we have

$$|c_j |x - x_j|| \leq c_j(|x| + |x_j|) \leq \frac{R+1}{2^j}.$$

By the Weierstrass M-test (Theorem 3.5.15 below) and Example 2.4.2–1 the lemma follows. ▼

It follows immediately from the lemma that the series defining f converges pointwise, and so f is well-defined, and is moreover convex by Theorem 3.5.26. Now we show that f is differentiable at points in $\mathbb{R} \setminus C$.

2 Lemma *The function* f *is differentiable at every point in* $\mathbb{R} \setminus$ C*.*

Proof Let us denote $g_j(x) = c_j|x - x_j|$. Let $x_0 \in \mathbb{R} \setminus \mathrm{C}$ and define, for each $j \in \mathbb{Z}_{>0}$,

$$h_{j,x_0} = \begin{cases} \frac{g_j(x) - g_j(x_0)}{x - x_0}, & x \neq x_0, \\ g_j'(x_0), & x = x_0, \end{cases}$$

noting that the functions g_j, $j \in \mathbb{Z}_{>0}$, are differentiable at points in $\mathbb{R} \setminus C$.

Let $j \in \mathbb{Z}$. We claim that if $x_0 \neq x_j$ then

$$|h_{j,x_0}(x)| \leq \frac{3}{2^j} \tag{3.10}$$

for all $x \in \mathbb{R}$. We consider three cases.

(a) $x = x_0$: Note that g_j is differentiable at $x = x_0$ and that $|g_j'(x_0)| = c_j \leq \frac{1}{2^j} < \frac{3}{2^j}$. Thus the estimate (3.10) holds when $x = x_0$.

(b) $x \neq x_0$ and $(x - x_j)(x_0 - x_j) > 0$: We have

$$|h_{j,x_0}(x)| = c_j\left|\frac{(x - x_j) - (x_0 - x_j)}{x - x_0}\right| = a_j \leq \frac{1}{2^j} < \frac{3}{2^j},$$

giving (3.10) in this case.

(c) $x \neq x_0$ and $(x - x_j)(x_0 - x_j) < 0$: We have

$$|h_{j,x_0}(x)| = c_j\left|\frac{(x - x_j) - (x_j - x_0)}{x - x_0}\right| = c_j\left|1 + \frac{2(x_0 - x_j)}{x_0 - x}\right| \leq \frac{1}{2^j}\left|1 + \frac{2(x_0 - x_j)}{x_0 - x}\right|.$$

Since $(x - x_j)$ and $x_0 - x_j$ have opposite sign, this implies that either (1) $x < x_j$ and $x_0 > x_j$ or (2) $x > x_j$ and $x_0 < x_j$. In either case, $|x_0 - x_j| < |x_0 - x|$. This, combined with our estimate above, gives (3.10) in this case.

Now, given (3.10), we can use the Weierstrass M-test (Theorem 3.5.15 below) and Example 2.4.2–1 to conclude that $\sum_{j=1}^{\infty} h_{j,x_0}$ converges uniformly on $\mathbb{R}$ for each $x_0 \in \mathbb{R} \setminus C$.

Now we prove that f is differentiable at $x_0 \in \mathbb{R} \setminus C$. If $x \neq x_0$ then the definition of the functions h_{j,x_0}, $j \in \mathbb{Z}_{>0}$, gives

$$\frac{f(x) - f(x_0)}{x - x_0} = \sum_{j=1}^{\infty} h_{j,x_0}(x),$$

the latter sum making sense since we have shown that it converges uniformly. Moreover, since the functions g_j, $j \in \mathbb{Z}_{>0}$, are differentiable at x_0, it follows that, for each $j \in \mathbb{Z}_{>0}$,

$$\lim_{x \to x_0} h_{j,x_0}(x) = \lim_{x \to x_0} \frac{g_j(x) - g_j(x_0)}{x - x_0} = g_j'(x_0) = h_{j,x_0}(x_0).$$

That is, h_{j,x_0} is continuous at x_0. It is clear that h_{j,x_0} is continuous at all $x \neq x_0$. Thus, since $\sum_{j=1}^{\infty} h_{j,x_0}$ converges uniformly, the limit function is continuous by Theorem 3.5.8. Thus we have

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} \sum_{j=1}^{\infty} h_{j,x_0}(x) = \sum_{j=1}^{\infty} h_{j,x_0}(x_0) = \sum_{j=1}^{\infty} g_j'(x_0).$$

This gives the desired differentiability since the last series converges. ▼

Finally, we show that f is not differentiable at points in C.

3 Lemma *The function* f *is not differentiable at every point in* C.

Proof For $k \in \mathbb{Z}_{>0}$, let us write

$$f(x) = g_k(x) + \underbrace{\sum_{\substack{j=1\\ j\neq k}} g_j(x)}_{f_j(x)}.$$

The arguments from the proof of the preceding lemma can be applied to show that the function f_j defined by the sum on the right is differentiable at x_k. Since g_k is not differentiable at x_k, we conclude that f cannot be differentiable at x_k by Proposition 3.2.10. ▼

This shows that the conclusions of Proposition 3.2.29(iv) cannot generally be improved. •

3.2.7 Piecewise differentiable functions

In Section 3.1.7 we considered functions that were piecewise continuous. In this section we consider a class of piecewise continuous functions that have additional properties concerning their differentiability. We let $I \subseteq \mathbb{R}$ be an interval with $f\colon I \to \mathbb{R}$ a function. In Section 3.1.7 we defined the notation $f(x-)$ and $f(x+)$. Here we also define

$$f'(x-) = \lim_{\epsilon\downarrow 0} \frac{f(x-\epsilon) - f(x-)}{-\epsilon}, \quad f'(x+) = \lim_{\epsilon\downarrow 0} \frac{f(x+\epsilon) - f(x+)}{\epsilon}.$$

These limits, of course, may fail to exist, or even to make sense if $x \in \mathrm{bd}(I)$.

Now, recalling the notion of a partition from Definition 2.5.7, we make the following definition.

3.2.32 Definition (Piecewise differentiable function) A function $f\colon [a,b] \to \mathbb{R}$ is ***piecewise differentiable*** if there exists a partition $P = (I_1, \dots, I_k)$, with $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$, of $[a,b]$ with the following properties:

(i) $f|\,\mathrm{int}(I_j)$ is differentiable for each $j \in \{1, \dots, k\}$;

(ii) for $j \in \{1, \dots, k-1\}$, the limits $f(x_j+)$, $f(x_j-)$, $f'(x_j+)$, and $f'(x_j-)$ exist;

(iii) the limits $f(a+)$, $f(b-)$, $f'(a+)$, and $f'(b-)$ exist. •

It is evident that a piecewise differentiable function is piecewise continuous. It is not surprising that the converse is not true, and a simple example of this will be given in the following collection of examples.

3.2.33 Examples (Piecewise differentiable functions)

1. Let $I = [-1, 1]$ and define $f\colon I \to \mathbb{R}$ by
$$f(x) = \begin{cases} 1 + x, & x \in [-1, 0], \\ 1 - x, & (0, 1]. \end{cases}$$
One verifies that f is differentiable on $(-1, 0)$ and $(0, 1)$. Moreover, we compute the limits
$$\begin{gathered} f(-1+) = 0, \quad f'(-1+) = 1, \quad f(1-) = 0, \quad f'(1-) = -1, \\ f(0-) = 1, \quad f(0+) = 1, \quad f'(0-) = 1, \quad f'(0+) = -1. \end{gathered}$$
Thus f is piecewise differentiable. Note that f is also continuous.
2. Let $I = [-1, 1]$ and define $f\colon I \to \mathbb{R}$ by $f(x) = \operatorname{sign}(x)$. On $(-1, 0)$ and $(0, 1)$ we note that f is differentiable. Moreover, we compute
$$\begin{gathered} f(-1+) = -1, \quad f'(-1+) = 0, \quad f(1-) = 1, \quad f'(1-) = 0, \\ f(0-) = -1, \quad f(0+) = 1, \quad f'(0-) = 0, \quad f'(0+) = 0. \end{gathered}$$
Note that it is important here to *not* compute the limits $f'(0-)$ and $f'(0+)$ using the formulae
$$\lim_{\epsilon \downarrow 0} \frac{f(0 - \epsilon) - f(0)}{-\epsilon}, \quad \lim_{\epsilon \downarrow 0} \frac{f(0 + \epsilon) - f(0)}{\epsilon}.$$
Indeed, these limits do not exist, where as the limits $f'(0-)$ and $f'(0+)$ do exist. In any event, f is piecewise differentiable, although it is not continuous.
3. Let $I = [0, 1]$ and define $f\colon I \to \mathbb{R}$ by $f(x) = \sqrt{x(1 - x)}$. On $(0, 1)$, f is differentiable. Also, the limits $f(0+)$ and $f(1-)$ exist. However, the limits $f'(0+)$ and $f'(1-)$ do not exist, as we saw in Example 3.2.3–3. Thus f is not piecewise differentiable. However, it is continuous, and therefore piecewise continuous, on $[0, 1]$. •

3.2.8 Notes

It was Weierstrass who first proved the existence of a continuous but nowhere differentiable function. The example Weierstrass gave was
$$\tilde{f}(x) = \sum_{j=0}^{\infty} b^n \cos(a^n \pi x),$$
where $b \in (0, 1)$ and a satisfies $ab > \frac{3}{2}\pi + 1$. It requires a little work to show that this function is nowhere differentiable. The example we give as Example 3.2.9 is fairly simple by comparison, and is taken from the paper of **JM:53**.

Example 3.2.31–2 if from **[SS/EES:04]**

Exercises

3.2.1 Let $I \subseteq \mathbb{R}$ be an interval and let $f, g\colon I \to \mathbb{R}$ be differentiable. Is it true that the functions

$$I \ni x \mapsto \min\{f(x), g(x)\} \in \mathbb{R}, \qquad I \ni x \mapsto \max\{f(x), g(x)\} \in \mathbb{R},$$

are differentiable? If it is true provide a proof, if it is not true, give a counterexample.

Section 3.3

ℝ-valued functions of bounded variation

In this section we present a class of functions, functions of so-called bounded variation, that are larger than the set of differentiable functions. However, they are sufficiently friendly that they often play a distinguished rôle in certain parts of signal theory, as evidenced by the theorems of Jordan concerning inversion of Fourier transforms (see Theorems 12.2.31 and 13.2.24). It is often not obvious after an initial reading on the topic of functions of bounded variation, just why such functions are important. Historically, the class of functions of bounded variation arose out of the desire to understand functions that are sums of functions that are monotonically increasing (see Definition 3.1.27 for the definition). Indeed, as we shall see in Theorem 3.3.3, functions of bounded variation and monotonically increasing functions are inextricably linked. The question about the importance of functions of bounded variation can thus be reduced to the question about the importance of monotonically increasing functions. An intuitive reason why such functions might be interesting is that many of the functions one encounters in practice, while not themselves increasing or decreasing, have intervals on which they *are* increasing or decreasing. Thus one hopes that, by understanding increasing or decreasing functions, one can understand more general functions.

It is also worth mentioning here that the class of functions of bounded variation arise in functional analysis as the topological dual to Banach spaces of continuous functions. In this regard, we refer the reader to Theorem **??**.***missing stuff***

Do I need to read this section? This section should be strongly considered for omission on a first read, and then referred to when the concept of bounded variation comes up in subsequent chapters, namely in Chapters 12 and 13. Such an omission is suggested, not because the material is unimportant or uninteresting, but rather because it constitutes a significant diversion that might be better left until it is needed. •

3.3.1 Functions of bounded variation on compact intervals

In this section we define functions of bounded variation on intervals that are compact. In the next section we shall extend these ideas to general intervals. For a compact interval I, recall that $\mathrm{Part}(I)$ denotes the set of partitions of I, and that if $P \in \mathrm{Part}(I)$ then $\mathrm{EP}(P)$ denotes the endpoints of the intervals comprising P (see the discussion surrounding Definition 2.5.7).

3.3.1 Definition (Total variation, function of bounded variation) For $I = [a, b]$ a compact interval and $f\colon I \to \mathbb{R}$ a function on I, the ***total variation*** of f is given by

$$\mathrm{TV}(f) = \sup\Big\{\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| \;\Big|\; (x_0, x_1, \dots, x_k) = \mathrm{EP}(P),\ P \in \mathrm{Part}([a, b])\Big\}.$$

If $\mathrm{TV}(f) < \infty$ then f has ***bounded variation***. •

Let us characterise real functions of bounded variation on compact intervals. The principal part of this characterisation is the decomposition of a function of bounded variation into the difference of monotonically increasing functions. However, another interesting characterisation involves the following idea which relies on the notion of the graph of a function, introduced following Definition 1.3.1.

3.3.2 Definition (Arclength of the graph of a function) Let $[a, b]$ be a compact interval and let $f\colon [a, b] \to \mathbb{R}$ be a function. The ***arclength*** of $\mathrm{graph}(f)$ is defined to be

$$\ell(\mathrm{graph}(f)) = \sup\Big\{\sum_{j=1}^{k}\Big((f(x_j) - f(x_{j-1}))^2 + (x_j - x_{j-1})^2\Big)^{1/2}\,\Big|\; (x_0, x_1, \dots, x_k) = \mathrm{EP}(P),\ P \in \mathrm{Part}([a, b])\Big\}.$$ •

We now have the following result which characterises functions of bounded variation.

3.3.3 Theorem (Characterisation of functions of bounded variation) *For a compact interval* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *and a function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$*, the following statements are equivalent:*

(i) f *has bounded variation;*

(ii) there exists monotonically increasing functions $\mathrm{f}_+, \mathrm{f}_-\colon \mathrm{I} \to \mathbb{R}$ *such that* $\mathrm{f} = \mathrm{f}_+ - \mathrm{f}_-$ *(**Jordan**[8] **decomposition** of a function of bounded variation);*

(iii) the graph of f *has finite arclength in* $\mathbb{R}^2$*.*

Furthermore, each of the preceding three statements implies the following:

(iv) the following limits exist:

(a) $\mathrm{f}(\mathrm{a}+)$*;*

(b) $\mathrm{f}(\mathrm{b}-)$*;*

(c) $\mathrm{f}(\mathrm{x}+)$ *and* $\mathrm{f}(\mathrm{x}-)$ *for all* $\mathrm{x} \in \mathrm{int}(\mathrm{I})$*,*

(v) f *is continuous except at a countable number of points in* I*,*

(vi) f *possesses a derivative almost everywhere in* I*.*

Proof (i) $\implies$ (ii) Define $\mathrm{V}(f)(x) = \mathrm{TV}(f|[a, x])$ so that $x \mapsto \mathrm{V}(f)(x)$ is a monotonic function. Let us define

$$f_+(x) = \tfrac{1}{2}(\mathrm{V}(f)(x) + f(x)), \quad f_-(x) = \tfrac{1}{2}(\mathrm{V}(f)(x) - f(x)). \tag{3.11}$$

Since we obviously have $f = f_+ - f_-$, this part of the theorem will follow if f_+ and f_- can be shown to be monotonic. Let $\xi_2 > \xi_1$ and let $(x_0, x_1, \dots, x_k)$ be the endpoints of

[8] Marie Ennemond Camille Jordan (1838–1922) was a French mathematician who made significant contributions to the areas of algebra, analysis, complex analysis, and topology. He wrote a three volume treatise on analysis entitled *Cours d'analyse de l'École Polytechnique* which was quite influential.

a partition of $[a, \xi_1]$. Then $(x_0, x_1, \dots, x_k, x_{k+1} = \xi_2)$ are the endpoints of a partition of $[a, \xi_2]$. We have the inequalities

$$\mathrm{V}(f)(\xi_2) \geq \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| + |f(\xi_2) - f(\xi_1)|.$$

Since this is true for any partition of $[a, \xi_1]$ we have

$$\mathrm{V}(f)(\xi_2) \geq \mathrm{V}(f)(\xi_1) + |f(\xi_2) - f(\xi_1)|.$$

We then have

$$\begin{aligned} 2f_+(\xi_2) &= \mathrm{V}(f)(\xi_2) + f(\xi_2) \\ &\geq \mathrm{V}(f)(\xi_1) + f(\xi_1) + |f(\xi_2) - f(\xi_1)| + f(\xi_2) - f(\xi_1) \\ &\geq \mathrm{V}(f)(\xi_1) + f(\xi_1) = 2f_+(\xi_1) \end{aligned}$$

and

$$\begin{aligned} 2f_-(\xi_2) &= \mathrm{V}(f)(\xi_2) - f(\xi_2) \\ &\geq \mathrm{V}(f)(\xi_1) - f(\xi_1) + |f(\xi_2) - f(\xi_1)| - f(\xi_2) + f(\xi_1) \\ &\geq \mathrm{V}(f)(\xi_1) - f(\xi_1) = 2f_+(\xi_1), \end{aligned}$$

giving this part of the theorem.

(ii) $\Longrightarrow$ (i) If f is monotonically increasing and if $(x_0, x_1, \dots, x_k)$ are the endpoints for a partition of $[a, b]$, then

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| = \sum_{j=1}^{k} (f(x_j) - f(x_{j-1})) = f(b) - f(a).$$

Thus monotonically increasing functions, and similarly monotonically decreasing functions, have bounded variation. Now consider two functions f and g, both of bounded variation. By part (i) of Proposition 3.3.12, $f + g$ is also of bounded variation. In particular, the sum of a monotonically increasing and a monotonically decreasing function will be a function of bounded variation.

(i) $\Longleftrightarrow$ (iii) First we note that, for any $a, b \in \mathbb{R}$,

$$(|a| + |b|)^2 = a^2 + b^2 + 2|a||b|,$$

from which we conclude that $(a^2 + b^2)^{1/2} \leq |a| + |b|$. Therefore, if $(x_0, x_1, \dots, x_k)$ are the endpoints of a partition of $[a, b]$, then

$$\begin{aligned} \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| &\leq \sum_{j=1}^{k} \Big((f(x_j) - f(x_{j-1}))^2 + (x_j - x_{j-1})^2 \Big)^{1/2} \\ &\leq \sum_{j=1}^{k} \Big(|f(x_j) - f(x_{j-1})| + |x_j - x_{j-1}| \Big) = \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| + b - a. \quad (3.12) \end{aligned}$$

This implies that

$$\mathrm{TV}(f) \leq \ell(\mathrm{graph}(f)) \leq \mathrm{TV}(f) + b - a,$$

from which this part of the result follows.

(iv) Let f_+ and f_- be monotonically increasing functions as per part (ii). By Theorem 3.1.28 we know that the limits asserted in this part of the theorem hold for both f_+ and f_-. This part of the theorem now follows from Propositions 2.3.23 and 2.3.29.

(v) This follows from Theorem 3.1.28 and Proposition 3.1.15, using the decomposition $f = f_+ - f_-$ from part (ii).

(vi) Again using the decomposition $f = f_+ - f_-$ from part (ii), this part of the theorem follows from Theorem 3.2.26 and Proposition 3.2.10. ■

3.3.4 Remark We comment the converses of parts (iv), (v), and (vi) of Theorem 3.3.3 do not generally hold. This is because, as we shall see in Example 3.3.5–4, continuous functions are not necessarily of bounded variation. •

Let us give some examples of functions that have and do not have bounded variation.

3.3.5 Examples (Functions of bounded variation on compact intervals)

1. On $[0, 1]$ define $f\colon [0, 1] \to \mathbb{R}$ by $f(x) = c$, for $c \in \mathbb{R}$. We easily see that $\mathrm{TV}(f) = 0$, so f has bounded variation.
2. On $[0, 1]$ consider the function $f\colon [0, 1] \to \mathbb{R}$ defined by $f(x) = x$. We claim that f has bounded variation. Indeed, if $(x_0, x_1, \dots, x_k)$ are the endpoints of a partition of $[0, 1]$, then we have

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| = \sum_{j=1}^{k} |x_j - x_{j-1}| = 1 - 0 = 1,$$

thus giving f as having bounded variation.

Note that f is itself a monotonically increasing function, so that for part (ii) of Theorem 3.3.3 we may take $f_+ = f$ and f_- to be the zero function. However, we can also write $f = g_+ - g_-$ where $g_+(x) = 2x$ and $g_-(x) = x$. Thus the decomposition of part (ii) of Theorem 3.3.3 is not unique.

3. On $I = [0, 1]$ consider the function

$$f(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}] \\ -1, & x \in (\frac{1}{2}, 1]. \end{cases}$$

We claim that $\mathrm{TV}(f) = 1$. Let $(x_0, x_1, \dots, x_k)$ be the endpoints of a partition of $[0, 1]$. Let $\bar{k}$ be the least element in $\{1, \dots, k\}$ for which $x_{\bar{k}} > \frac{1}{2}$. Then we have

$$\begin{aligned} \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| = \sum_{j=1}^{\bar{k}-1} |f(x_j) - f(x_{j-1})| + \sum_{j=\bar{k}+1}^{k} |f(x_j) - f(x_{j-1})| \\ + |f(x_{\bar{k}}) - f(x_{\bar{k}-1})| = 1. \end{aligned}$$

This shows that $\mathrm{TV}(f) = 1$ and so f has bounded variation. Note that this also shows that functions of bounded variation need not be continuous. This, along with the next example, shows that the relationship between continuity and bounded variation is not a straightforward one.

4. Consider the function on $I = [0,1]$ defined by

$$f(x) = \begin{cases} x\sin\frac{1}{x}, & x \in (0,1], \\ 0, & x = 0. \end{cases}$$

We first claim that f is continuous. Clearly it is continuous at x provided that $x \neq 0$. To show continuity at $x = 0$, let $\epsilon \in \mathbb{R}_{>0}$ and note that, if $x < \epsilon$, we have $|f(x)| < \epsilon$, thus showing continuity.

However, f does not have bounded variation. Indeed, for $j \in \mathbb{Z}_{>0}$ denote $\xi_j = \frac{1}{(j+\frac{1}{2})\pi}$. Then, for $k \in \mathbb{Z}_{>0}$, consider the partition with endpoints

$$(x_0 = 0, x_1 = \xi_k, \dots, x_k + = \xi_1, x_{k+1} = 1).$$

Direct computation then gives

$$\begin{aligned}\sum_{j=1}^{k+1} |f(x_j) - f(x_{j-1})| &\geq \frac{2}{\pi}\sum_{j=1}^{k}\left|\frac{(-1)^j}{2j+1} - \frac{(-1)^{j-1}}{2j-1}\right| \\ &= \frac{2}{\pi}\sum_{j=1}^{k}\left|\frac{1}{2j+1} + \frac{1}{2j-1}\right| \geq \frac{2}{\pi}\sum_{j=1}^{k}\left|\frac{2}{2j+1}\right|.\end{aligned}$$

Thus

$$\mathrm{TV}(f) \geq \frac{2}{\pi}\sum_{j=1}^{\infty}\left|\frac{2}{2j+1}\right| = \infty,$$

showing that f has unbounded variation. •

3.3.2 Functions of bounded variation on general intervals

Now, with the definitions and properties of bounded variation for functions defined on compact intervals, we can sensibly define notions of variation for general intervals.

3.3.6 Definition (Bounded variation, locally bounded variation) Let I be an interval with $f\colon I \to \mathbb{R}$ a function.

(i) If $f|[a,b]$ is a function of bounded variation for every compact interval $[a,b] \subseteq I$, then f is a function of ***locally bounded variation***.

(ii) If $\sup\{\mathrm{TV}(f|[a,b]) \mid [a,b] \subseteq I\} < \infty$, then f is a function of ***bounded variation***. •

3.3.7 Remark (Properties of functions of locally bounded variation) We comment that the characterisations of functions of bounded variation given in Theorem 3.3.3 carry over to functions of locally bounded variation in the sense that the following statements are equivalent for a function $f\colon I \to \mathbb{R}$ defined on a general interval I:

1. f has locally bounded variation;

2. there exists monotonically increasing functions $f_+, f_- \colon I \to \mathbb{R}$ such that $f = f_+ - f_-$.

Furthermore, each of the preceding two statements implies the following:

3. the following limits exist:
   (a) $f(a+)$;
   (b) $f(b-)$;
   (c) $f(x+)$ and $f(x-)$ for all $x \in \operatorname{int}(I)$,
4. f is continuous except at a countable number of points in I,
5. f possesses a derivative almost everywhere in I.

These facts follow easily from the definition of locally bounded variation, along with facts about countable sets, and sets of measure zero. We leave the details to the reader as Exercise 3.3.4. •

3.3.8 Notation ("Locally bounded variation" versus "bounded variation") These extended definitions agree with the previous ones in that, when I is compact, (1) the new definition of a function of bounded variation agrees with that of Definition 3.3.1 and (2) the definition of a function of bounded variation agrees with the definition of a function of locally bounded variation. The second point is particularly important to remember, because most of the results in the remainder of this section will be stated for functions of locally bounded variation. Our observation here is that these results automatically apply to functions of bounded variation, as per Definition 3.3.1. For this reason, we will generally default from now on to using "locally bounded variation" in place of "bounded variation," reserving the latter for when it is intended in its distinct place when the interval of definition of a function is compact. •

Let us give some examples of functions that do and no not have locally bounded variation.

3.3.9 Examples (Functions of locally bounded variation on general intervals)

1. Let $I \subseteq \mathbb{R}$ be an arbitrary interval, let $c \in \mathbb{R}$, and consider the function $f \colon I \to \mathbb{R}$ defined by $f(x) = c$. Applying the definition shows that $\mathrm{TV}(f|[a,b])(x) = 0$ for all compact intervals $[a,b] \subseteq I$, no matter the character of I. Thus constant functions, unsurprisingly, have locally bounded variation.
2. Let us consider the function $f \colon I \to \mathbb{R}$ on $I = [0,\infty)$ defined by $f(x) = x$. We claim that f has locally bounded variation. Indeed, let $[a,b] \subseteq I$ and consider a partition of $[a,b]$ with endpoints $(x_0, x_1, \dots, x_k)$. We have

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| = \sum_{j=1}^{k} (x_j - x_{j-1}) = b - a.$$

This shows that f has locally bounded variation. However, since $b - a$ can be arbitrarily large, f does not have bounded variation.

3. On the interval $I = (0,1]$ consider the function $f\colon I \to \mathbb{R}$ defined by $f(x) = \frac{1}{x}$. Note that, for $[a,b] \subseteq (0,1]$, the function $f|[a,b]$ is monotonically decreasing, and so has bounded variation. We can thus conclude that f is a function of locally bounded variation. We claim that f does not have bounded variation. To see this, note that if $(x_0, x_1, \dots, x_k)$ are the endpoints of a partition of $[a,b] \subseteq (0,1]$, then it is easy to see that, since f is strictly monotonically decreasing and continuous that $(f(x_k), \dots, f(x_1), f(x_0))$ are the endpoints of a partition of $[f(x_k), f(x_0)]$. We thus have

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| = f(x_0) - f(x_k).$$

Since $f(x_0)$ can be made arbitrarily large by choosing a small, it follows that f cannot have bounded variation. •

We close this section by introducing the notion of the variation of a function, and giving a useful property of this concept.

3.3.10 Definition (Variation of a function of bounded variation) Let $I \subseteq \mathbb{R}$ be an interval, let $a \in I$, let $f\colon I \to \mathbb{R}$ be a function of locally bounded variation, and define $\mathrm{V}_a(f)\colon I \to \mathbb{R}_{>0}$ by

$$\mathrm{V}_a(f)(x) = \begin{cases} \mathrm{TV}(f|[x,a]), & x < a, \\ 0, & x = a, \\ \mathrm{TV}(f|[a,x]), & x > a. \end{cases}$$

The function $\mathrm{V}_a(f)$ is the ***variation*** of f with reference point a. •

One can easily check that the choice of a in the definition of $\mathrm{V}_a(f)$ serves only to shift the values of the function. Thus the essential features of the variation are independent of the reference point.

When a function of bounded variation is continuous, so too is its variation.

3.3.11 Proposition (The variation of a continuous function is continuous and vice versa) *Let* $I \subseteq \mathbb{R}$ *be an interval, let* $a \in I$, *and let* $f\colon I \to \mathbb{R}$ *be a function of locally bounded variation. Then* f *is continuous at* $x \in I$ *if and only if* $\mathrm{V}_a(f)$ *is continuous at* x. *Moreover, if* f *is a continuous function of bounded variation, then* $f = f_+ - f_-$ *where* f_+ *and* f_- *are continuous monotonically increasing functions.*

Proof The general result follows easily from the case where $I = [a,b]$ is compact. Furthermore, in this case it suffices to consider the variation of f with reference points a or b. We shall consider only the reference point a, since the other case follows in much the same manner.

Suppose that f is continuous at $x_0 \in I$ and let $\epsilon \in \mathbb{R}_{>0}$. First suppose that $x_0 \in [a,b)$, and let $\delta \in \mathbb{R}_{>0}$ be chosen such that $x \in \mathsf{B}(\delta, x_0) \cap I$ implies that $|f(x) - f(x_0)| < \frac{\epsilon}{2}$. Choose a partition of $[x_0, b]$ with endpoints $(x_0, x_1, \dots, x_k)$ such that

$$\mathrm{TV}(f|[x_0,b]) - \tfrac{\epsilon}{2} \le \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})|. \tag{3.13}$$

We may without loss of generality suppose that $x_1 - x_0 < \delta$. Indeed, if this is not the case, we may add a new endpoint to our partition, noting that the estimate (3.13) will hold for the new partition. We then have

$$\begin{aligned}\mathrm{TV}(f|[x_0,b]) - \tfrac{\epsilon}{2} &\le |f(x_1) - f(x_0)| + \sum_{j=2}^{k} |f(x_j) - f(x_{j-1})| \\ &\le \tfrac{\epsilon}{2} + \sum_{j=2}^{k} |f(x_j) - f(x_{j-1})| \le \tfrac{\epsilon}{2} + \mathrm{TV}(f|[x_1,b]).\end{aligned}$$

This then gives

$$\mathrm{TV}(f|[x_0,b]) - \mathrm{TV}(f|[x_1,b]) = \mathrm{V}_a(f)(x_1) - \mathrm{V}_a(f)(x_0) < \epsilon.$$

Since this holds for any partition for which $x_1 - x_0 < \delta$, it follows that $\lim_{x \downarrow x_0} \mathrm{V}_a(f)(x) = \mathrm{V}_a(f)(x_0)$ for every $x_0 \in [a,b)$ at which f is continuous. One can similarly show that $\lim_{x \uparrow x_0} \mathrm{V}_a(f)(x) = \mathrm{V}_a(f)(x_0)$ for every $x_0 \in (a,b]$ at which f is continuous. This gives the result by Theorem 3.1.3.

Suppose that $\mathrm{V}_a(f)$ is continuous at $x_0 \in I$ and let $\epsilon \in \mathbb{R}_{>0}$. Choose $\delta \in \mathbb{R}_{>0}$ such that $|\mathrm{V}_a(f)(x) - \mathrm{V}_a(f)(x_0)| < \epsilon$ for $x \in \overline{\mathsf{B}}(2\delta, x_0)$. Then, for $x \in \overline{\mathsf{B}}(2\delta, x_0)$ with $x > x_0$,

$$|f(x) - f(x_0)| \le \mathrm{TV}(f|[x_0,x]) = \mathrm{V}_a(f)(x) - \mathrm{V}_a(f)(x_0) < \epsilon,$$

using the fact that (x_0, x) are the endpoints of a partition of $[x_0, x]$. In like manner, if $x \in \overline{\mathsf{B}}(2\delta, x_0)$ with $x > x_0$, then

$$|f(x) - f(x_0)| \le \mathrm{TV}(f|[x,x_0]) = \mathrm{V}_a(f)(x_0) - \mathrm{V}_a(f)(x) < \epsilon.$$

Thus $|f(x) - f(x_0)| < \epsilon$ for every $x \in \overline{\mathsf{B}}(2\delta, x_0)$, and so for every $x \in \mathsf{B}(\delta, x_0)$, giving continuity of f at x_0.

The final assertion follows from the definition of the Jordan decomposition given in (3.11). ■

3.3.3 Bounded variation and operations on functions

In this section we illustrate how functions of locally bounded variation interact with the usual operations one performs on functions.

3.3.12 Proposition (Addition and multiplication, and locally bounded variation) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval and let* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *be functions of locally bounded variation. Then the following statements hold:*

(i) $\mathrm{f} + \mathrm{g}$ *is a function of locally bounded variation;*

(ii) fg *is a function of locally bounded variation;*

(iii) if additionally there exists $\alpha \in \mathbb{R}_{>0}$ *such that* $|\mathrm{g}(\mathrm{x})| \ge \alpha$ *for all* $\mathrm{x} \in \mathrm{I}$, *then* $\frac{\mathrm{f}}{\mathrm{g}}$ *is a function of locally bounded variation.*

Proof Without loss of generality we may suppose that $I = [a,b]$ is a compact interval.

(i) Let $(x_0, x_1, \dots, x_k)$ be the endpoints for a partition of $[a, b]$ and compute

$$\sum_{j=1}^{k} |f(x_j) + g(x_j) - f(x_{j-1}) - g(x_{j-1})| \le \sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| + \sum_{j=1}^{k} |g(x_j) - g(x_{j-1})|$$

using the triangle inequality. It then follows from Proposition 2.2.27 that $\mathrm{TV}(f + g) \le \mathrm{TV}(f) + \mathrm{TV}(g)$, and so $f + g$ has locally bounded variation.

(ii) Let

$$M_f = \sup\{|f(x)| \mid x \in [a, b]\}, \quad M_g = \sup\{|g(x)| \mid x \in [a, b]\}.$$

Then, for a partition of $[a, b]$ with endpoints $(x_0, x_1, \dots, x_k)$, compute

$$\begin{aligned}\sum_{j=1}^{k} |f(x_j)g(x_j) - f(x_{j-1})g(x_{j-1})| &\le \sum_{j=1}^{k} |f(x_j)g(x_j) - f(x_{j-1})g(x_j)| \\ &\quad + \sum_{j=1}^{k} |f(x_{j-1})g(x_j) - f(x_{j-1})g(x_{j-1})| \\ &\le \sum_{j=1}^{k} M_g |f(x_j) - f(x_{j-1})| + \sum_{j=1}^{k} M_f |g(x_j) - g(x_{j-1})| \\ &\le M_g \,\mathrm{TV}(f) + M_f \,\mathrm{TV}(g),\end{aligned}$$

giving the result.

(iii) Let $(x_0, x_1, \dots, x_k)$ be a partition of $[a, b]$ and compute

$$\sum_{j=1}^{k} \left| \frac{1}{g(x_j)} - \frac{1}{g(x_{j-1})} \right| = \sum_{j=1}^{k} \left| \frac{g(x_{j-1}) - g(x_j)}{g(x_j)g(x_{j-1})} \right| \le \sum_{j=1}^{k} \left| \frac{g(x_j) - g(x_{j-1})}{\alpha^2} \right| \le \frac{\mathrm{TV}(g)}{\alpha^2}.$$

Thus $\frac{1}{g}$ has locally bounded variation, and this part of the result follows from part (ii). ■

Next we show that to determine whether a function has locally bounded variation, one can break up the interval of definition into subintervals.

3.3.13 Proposition (Locally bounded variation on disjoint subintervals) *Let $I \subseteq \mathbb{R}$ be an interval and let $I = I_1 \cup I_2$, where $I_1 \cap I_2 = \{c\}$, where c is the right endpoint of I_1 and the left endpoint of I_2. Then $f\colon I \to \mathbb{R}$ has locally bounded variation if and only if $f|I_1$ and $f|I_2$ have locally bounded variation.*

Proof It suffices to consider the case where $I = [a, b]$, $I_1 = [a, c]$, and $I_2 = [c, b]$. First let $(x_0, x_1, \dots, x_k)$ be the endpoints of a partition of $[a, c]$ and let $(y_0, y_1, \dots, y_l)$ be the endpoints of a partition of $[c, b]$. Then

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| + \sum_{j=1}^{l} |f(y_j) - f(y_{j-1})| \le \mathrm{TV}(f),$$

which shows that $\mathrm{TV}(f|[a, c]) + \mathrm{TV}(f|[c, b]) \le \mathrm{TV}(f)$. Now let $(x_0, x_1, \dots, x_k)$ be the endpoints of a partition of $[a, b]$. If c is not one of the endpoints, then let $m \in \{1, \dots, k-1\}$ satisfy $x_{m-1} < c < x_m$, and define a new partition with endpoints

$$(y_0 = x_0, y_1 = x_1, \dots, ym - 1 = x_{m-1}, y_m = c, y_{m+1} = x_m, \dots, y_{k+1} = x_k).$$

Then

$$\begin{aligned}\sum_{j=1}^{k}|f(x_j)-f(x_{j-1})| &\leq \sum_{j=1}^{k+1}|f(y_j)-f(y_{j-1})|\\ &\leq \sum_{j=1}^{m}|f(y_j)-f(y_{j-1})| + \sum_{j=m}^{m+1}|f(y_j)-f(y_{j-1})|\\ &\leq \mathrm{TV}([a,c]) + \mathrm{TV}(f|[c,b]).\end{aligned}$$

This shows that $\mathrm{TV}(f) \leq \mathrm{TV}(f|[a,c]) + \mathrm{TV}(f|[c,b])$, which gives the result when combined with our previous estimate $\mathrm{TV}(f|[a,c]) + \mathrm{TV}(f|[c,b]) \leq \mathrm{TV}(f)$. ■

While Examples Example 3.3.5–3 and 4 illustrate that functions of locally bounded variation need not be continuous, and that continuous functions need not have locally bounded variation, the story for differentiability is more pleasant.

3.3.14 Proposition (Differentiable functions have locally bounded variation) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if the function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is differentiable with the derivative* f' *being locally bounded, then* f *has locally bounded variation. In particular, if* f *is of class* C^1*, then* f *is of locally bounded variation.*

Proof The general result follows from the case where $I = [a,b]$, so we suppose in the proof that I is compact. Let $(x_0, x_1, \dots, x_k)$ be a partition of $[a,b]$. By the Mean Value Theorem, for each $j \in \{1,\dots,k\}$ there exists $y_j \in (x_{j-1}, x_j)$ such that

$$f(x_j) - f(x_{j-1}) = f'(y_j)(x_j - x_{j-1}).$$

Moreover, since f' is bounded, let $M \in \mathbb{R}_{>0}$ satisfy $|f'(x)| < M$ for each $x \in [a,b]$. Then

$$\sum_{j=1}^{k}|f(x_j)-f(x_{j-1})| = \sum_{j=1}^{k}|f'(y_j)||x_j - x_{j-1}| \leq \sum_{j=1}^{k} M|x_j - x_{j-1}| = M(b-a).$$

The final assertion follows since, if f is of class C^1, then f' is continuous and so bounded by Theorem 3.1.22. ■

In the preceding result we asked that the derivative be locally bounded. This condition is essential, as the following example shows.

3.3.15 Example (A differentiable function that does not have bounded variation) We take $f\colon [-1,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2\sin(\frac{1}{x^2}), & x \neq 0,\\ 0, & x = 0.\end{cases}$$

We will show that this function is differentiable but does not have bounded variation. The differentiability of f at $x \neq 0$ follows from the product rule and the Chain

Rule since the functions $x \mapsto x^2$, $x \mapsto \frac{1}{x^2}$, and sin are all differentiable away from zero. Indeed, by the product and Chain Rule we have

$$f'(x) = 2x\sin(\tfrac{1}{x^2}) - \tfrac{2}{x}\cos(\tfrac{1}{x^2}).$$

For differentiability at $x = 0$ we compute

$$\lim_{h\to 0}\frac{f(0+h)-f(0)}{h} = \lim_{h\to 0}\frac{h^2\sin(\frac{1}{h^2})-0}{h} = \lim_{h\to 0} h\sin(\tfrac{1}{h^2}) = 0,$$

giving the derivative at $x = 0$ to be zero.

To show that f does not have bounded variation, for $j \in \mathbb{Z}_{>0}$ define

$$\xi_j = \frac{1}{\sqrt{(j+\frac{1}{2})\pi}}.$$

For $k \in \mathbb{Z}_{>0}$ define a partition of $[0,1]$ by asking that it have endpoints $(x_0, x_1 = \xi_k, \dots, x_k = \xi_1, x_{k+1})$. Then

$$\begin{aligned}\sum_{j=1}^{k+1}|f(x_j)-f(x_{j-1})| \ge \sum_{j=1}^{k}|f(x_j)-f(x_{j-1})| &= \frac{2}{\pi}\sum_{j=1}^{k}\Big|\frac{(-1)^j}{2j+1} - \frac{(-1)^{j-1}}{2j-1}\Big| \\ &\ge \frac{2}{\pi}\sum_{j=1}^{k}\Big|\frac{1}{2j+1}+\frac{1}{2j-1}\Big| \ge \frac{2}{\pi}\sum_{j=1}^{k}\Big|\frac{2}{2j+1}\Big|.\end{aligned}$$

Thus

$$\mathrm{TV}(f) \ge \frac{2}{\pi}\sum_{j=1}^{\infty}\Big|\frac{2}{2j+1}\Big| = \infty,$$

giving our assertion that f does not have bounded variation.

Note that it follows from Proposition 3.3.14 that f' is not bounded. This can be verified explicitly as well. •

While the composition of continuous functions is again a continuous function, and the composition of differentiable functions is again a differentiable function, the same assertion does not hold for functions of locally bounded variation.

3.3.16 Example (Compositions of functions of locally bounded variation need not be functions of locally bounded variation) Let $I = [-1,1]$ and define $f, g\colon I \to \mathbb{R}$ by $f(x) = x^{1/3}$ and

$$g(x) = \begin{cases} x^3(\sin\frac{1}{x})^3, & x \ne 0, \\ 0, & x = 0. \end{cases}$$

We claim that f and g are functions of bounded variation. To show that f has bounded variation, we note that f is monotonically increasing, and so necessarily of bounded variation by Theorem 3.3.3(ii). To show that g is of bounded variation, we shall show that it is of class C^1, and then use Proposition 3.3.14. Clearly g is differentiable with continuous derivative on the intervals $[-1,0)$ and $(0,1]$. Thus we need to show that g is differentiable at 0 with continuous derivative there. To see that g is differentiable at 0, we compute

$$\lim_{x\to 0}\frac{g(x)-g(0)}{x-0}=\lim_{x\to 0}x^2(\sin\tfrac{1}{x})^{1/3}=0,$$

since $\left|(\sin\frac{1}{x})^{1/3}\right|\le 1$. Thus $g'(0)=0$. We also can readily compute that $\lim_{x\downarrow 0}g'(x)=\lim_{x\uparrow 0}g'(x)=0$. Thus g' is also continuous at 0, so showing that g has bounded variation.

However, note that

$$f\circ g(x)=\begin{cases}x\sin\frac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

and in Example 3.3.5–4 we showed that this function does not have bounded variation on the interval $[0,1]$. Therefore, it cannot have bounded variation on the interval $[-1,1]$. This gives our desired conclusion that $f\circ g$ is not a function of bounded variation, even though both f and g are. •

3.3.4 Saltus functions

As we saw in part (v) of Theorem 3.3.3, a function of locally bounded variation is discontinuous at at most a countable set of points. Moreover, part (iv) of the same theorem indicates that all discontinuities are jump discontinuities. In the next section we shall see that it is possible to separate out these discontinuities into a single function which, when subtracted from a function of locally bounded variation, leaves a *continuous* function of locally bounded variation.

First we give a general definition, unrelated specifically to functions of locally bounded variation. For this definition we recall from Section 2.4.7 our discussion of sums over arbitrary index sets.

3.3.17 Definition (Saltus function) Let $I\subseteq\mathbb{R}$ be an interval and let I' be the interval obtained by removing the right endpoint from I, if I indeed contains its right endpoint; otherwise take $I'=I$. A ***saltus function***[9] on I is a function $j\colon I\to\mathbb{R}$ of the form

$$j(x)=\sum_{\xi\in(-\infty,x)\cap I}r_\xi+\sum_{\xi\in(-\infty,x]\cap I}l_\xi,$$

where $(r_\xi)_{\xi\in I'}$ and $(l_\xi)_{\xi\in I}$ are summable families of real numbers. •

This definition seems mildly ridiculous at a first read, in that there seems to be no reason why such a function should be of any interest. However, as we shall see, every function of locally bounded variation naturally gives rise to a saltus function. Before we get to this, let us look at some properties of saltus function. It might be

[9] "Saltus" is a Latin word meaning "to leap." Indeed, a saltus function is also frequently referred to as a ***jump function***.

helpful to note that the function of Example 3.2.28 is a saltus function, as is easily seen from its definition. Many of the general properties of saltus functions follow in the same manner as they did for that example.

3.3.18 Proposition (Continuity of saltus functions) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{j}\colon \mathrm{I} \to \mathbb{R}$ *is a saltus function given by*

$$\mathrm{j}(\mathrm{x}) = \sum_{\xi \in (-\infty,\mathrm{x}) \cap \mathrm{I}} \mathrm{r}_\xi + \sum_{\xi \in (-\infty,\mathrm{x}] \cap \mathrm{I}} \mathrm{l}_\xi,$$

then for $\mathrm{x} \in \mathrm{I}$ *the following statements are equivalent:*

(i) j *is continuous at* x*;*

(ii) $\mathrm{r}_\mathrm{x} = \mathrm{l}_\mathrm{x} = 0$*.*

Proof Let $\epsilon \in \mathbb{R}_{>0}$ and note that, as can be deduced from our proof of Proposition 2.4.33, there exists a finite set $A_\epsilon \subseteq I$ such that

$$\sum_{x \in I' \setminus A_\epsilon} |r_x| + \sum_{x \in I \setminus A_\epsilon} |l_x| \leq \epsilon,$$

where $I' = I \setminus \{b\}$ is I is an interval containing its right endpoint b, and $I' = I$ otherwise. Now, for $x \in I$, let $\delta \in \mathbb{R}_{>0}$ have the property that $\mathsf{B}(\delta, x) \cap A_\epsilon$ is either empty, or contains only x. For $y \in \mathsf{B}(\delta, x) \cap I$ with $y < x$ we have

$$|j(y) - j(x) - l_x| = \Big| \sum_{\xi \in [y,x)} r_\xi + \sum_{\xi \in [y,x)} l_\xi \Big| \leq \sum_{\xi \in I' \setminus A_\epsilon} |r_\xi| + \sum_{\xi \in I \setminus A_\epsilon} |l_\xi| < \epsilon.$$

Also, for $x < y$ we have

$$|j(y) - (j(x) + r_x)| = \Big| \sum_{\xi \in (x,y)} r_\xi + \sum_{\xi \in (x,y]} l_\xi \Big| \leq \sum_{\xi \in I' \setminus A_\epsilon} |r_\xi| + \sum_{\xi \in I \setminus A_\epsilon} |l_\xi| < \epsilon.$$

This gives $j(x-) = j(x) - l_x$ provided that x is not the left endpoint of I and $j(x+) = j(x) + r_x$ provided that x is not the right endpoint of I. Thus j is continuous at x if and only if $r_x = l_x = 0$. ■

3.3.19 Proposition (Saltus functions are of locally bounded variation) *If* I *is an interval and if* $\mathrm{j}\colon \mathrm{I} \to \mathbb{R}$ *is a saltus function, then* j *is a function of locally bounded variation.*

Proof We may without loss of generality suppose that $I = [a, b]$. Let us write

$$j(x) = \sum_{\xi \in (-\infty,x) \cap I} r_\xi + \sum_{\xi \in (-\infty,x] \cap I} l_\xi.$$

Let $x, y \in [a, b]$ with $x < y$. Then

$$j(y) - j(x) = r_x + l_y + \sum_{\xi \in (x,y)} (r_\xi + l_\xi).$$

Thus

$$|j(y) - j(x)| \leq \sum_{\xi \in [x,y)} |r_\xi| + \sum_{\xi \in (x,y]} |l_\xi|.$$

Now let $(x_0, x_1, \dots, x_m)$ be the endpoints of a partition of $[a, b]$. Then we compute

$$\sum_{k=1}^{m} |j(x_k) - j(x_{k-1})| \le \sum_{k=1}^{m} \Big(\sum_{\xi \in [x_{k-1}, x_k)} |u_\xi| + \sum_{\xi \in (x_{k-1}, x_k]} |l_\xi| \Big) \le \sum_{\xi \in [a,b)} |r_\xi| + \sum_{\xi \in (a,b]} |l_\xi|,$$

which gives the result. ■

Note then that we may now attribute to saltus functions all of the properties associated to functions of locally bounded variation, as presented in Theorem 3.3.3. In particular, a saltus function is differentiable almost everywhere. However, about the derivative of a saltus function, more can be said.

3.3.20 Proposition (Saltus functions have a.e. zero derivative) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{j}\colon \mathrm{I} \to \mathbb{R}$ *is a saltus function, then the set* $\{\mathrm{x} \in \mathrm{I} \mid \mathrm{j}'(\mathrm{x}) \neq 0\}$ *has measure zero.*

Proof Since j is of locally bounded variation, by Theorem 3.3.3(ii) we may write $j = j_+ - j_-$ for monotonically increasing functions j_+ and j_-. It then suffices to prove the result for the case when j is monotonically increasing, since the derivative is linear (Proposition 3.2.10) and since the union of two sets of measure zero is a set of measure zero (Exercise 2.5.9). As we saw in the proof of Proposition 3.3.18, $j(x-) = j(x) - l_x$ and $j(x+) = j(x) + r_x$. Therefore, if j is monotonically increasing, then $r_x \ge 0$ for all $x \in I'$ and $l_x \ge 0$ for all $x \in I$.

By Proposition 2.4.33 we may write

$$\{x \in I' \mid r_x \neq 0\} = \cup_{a \in A} \{\xi_a\}, \quad \{x \in I \mid l_x \neq 0\} = \cup_{b \in B} \{\eta_b\},$$

where the sets A and B are countable. For $x \in I$ define

$$A(x) = \{a \in A \mid \xi_a < x\}, \quad B(x) = \{b \in B \mid \eta_b \le x\}.$$

Then we have

$$\sum_{\xi \in (-\infty, x) \cap I} r_\xi = \sum_{a \in A(x)} r_{\xi_a}, \qquad \sum_{\xi \in (-\infty, x] \cap I} l_\xi = \sum_{b \in B(x)} r_{\eta_b}.$$

Now let us suppose that the sets A and B are well ordered and for $k \in \mathbb{Z}_{>0}$ define

$$A_k = \{a \in A \mid a \le k\}, \quad B_k = \{b \in B \mid b \le k\}$$

and

$$A_k(x) = \{a \in A_k \mid \xi_a < x\}, \quad B_k(x) = \{b \in B_k \mid \eta_b \le x\}.$$

We then define $j_k\colon I \to \mathbb{R}$ by

$$j_k(x) = \sum_{a \in A_k(x)} r_{\xi_a} + \sum_{b \in B_k(x)} r_{\eta_b}.$$

Now we use some facts from Section 3.5. Note the following facts:

1. for each $k \in \mathbb{Z}_{>0}$, the functions j_k are monotonically increasing since $r_x \ge 0$ for all $x \in I'$ and $l_x \ge 0$ for each $x \in I$;
2. for each $k \in \mathbb{Z}_{>0}$, the set $\{x \in I \mid j_k'(x) \neq 0\}$ is finite;
3. $\lim_{k \to \infty} j_k(x) = j(x)$ for each $x \in I$.

Therefore, we may apply Theorem 3.5.25 below to conclude that $j'(x) = 0$ almost everywhere. ■

3.3.21 **Remark (Functions with a.e. zero derivative need not be saltus functions)** Note that the Cantor function of Example 3.2.27 is a function with a derivative that is zero almost everywhere. However, since this function is continuous, it is not a saltus function. More precisely, according to Proposition 3.3.18, the Cantor function is a saltus function where the two families of summable numbers used to define it are both identically zero. That is to say, it is not an interesting saltus function. This observation will be important when we discuss the Lebesgue decomposition of a function of bounded variation in ***missing stuff*** •

3.3.5 The saltus function for a function of locally bounded variation

Now that we have outlined the general definition and properties of saltus functions, let us indicate how they arise from an attempt to generally characterise functions of locally bounded variation. Since functions of locally bounded variation are so tightly connected with monotonically increasing functions, we begin by constructing a saltus function associated to a monotonically increasing function.

3.3.22 **Proposition (Saltus function of a monotonically increasing function)** *Let* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *be a compact interval and let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be monotonically increasing. Define two families* $(\mathrm{r}_{\mathrm{f},\mathrm{x}})_{\mathrm{x}\in\mathrm{I}'}$ *and* $(\mathrm{l}_{\mathrm{f},\mathrm{x}})_{\mathrm{x}\in\mathrm{I}}$ *of real numbers by*

$$\begin{aligned}\mathrm{r}_{\mathrm{f},\mathrm{x}} &= \mathrm{f}(\mathrm{x}+) - \mathrm{f}(\mathrm{x}), \qquad \mathrm{x} \in [\mathrm{a}, \mathrm{b}),\\ \mathrm{l}_{\mathrm{f},\mathrm{a}} &= 0,\ \mathrm{l}_{\mathrm{f},\mathrm{x}} = \mathrm{f}(\mathrm{x}) - \mathrm{f}(\mathrm{x}-), \qquad \mathrm{x} \in (\mathrm{a}, \mathrm{b}],\end{aligned}$$

and let $\mathrm{j}_\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be defined by*

$$\mathrm{j}_\mathrm{f}(\mathrm{x}) = \sum_{\xi\in(-\infty,\mathrm{x})\cap\mathrm{I}} \mathrm{r}_{\mathrm{f},\xi} + \sum_{\xi\in(-\infty,\mathrm{x}]\cap\mathrm{I}} \mathrm{l}_{\mathrm{f},\xi}.$$

Then j_f *is a monotonically increasing saltus function, and the function* $\mathrm{f} - \mathrm{j}_\mathrm{f}$ *is a continuous monotonically increasing function.*

Proof Note that since f is monotonically increasing, $r_{f,x} \geq 0$ for all $x \in [a, b)$ and $l_{f,x} \geq 0$ for all $x \in [a, b]$. To show that j_f is a saltus function, it suffices to show that $(r_{f,x})_{x\in I'}$ and $(l_{f,x})_{x\in I}$ are summable. Let $(x_1, \dots, x_k)$ be a finite family of elements of $[a, b]$ (not necessarily the endpoints of a partition) and compute

$$\sum_{j=1}^{k}(r_{f,x_j} + l_{f,x_j}) = \sum_{j=1}^{k}(f(x_j+) - f(x_j-)) \leq f(b) - f(a).$$

Since this holds for every finite family $(x_1, \dots, x_k)$, we can assert that both families $(r_{f,x})_{x\in I'}$ and $(l_{f,x})_{x\in I}$ are summable.

Now let $x, y \in [a, b]$ with $x < y$. Take a partition of $[x, y]$ with endpoints $(x_0, x_1, \dots, x_k)$ and compute

$$\begin{aligned}&f(x+) - f(x) + \sum_{j=1}^{k}(f(x_j+) - f(x_j-)) + f(y) - f(y-),\\ &\qquad\qquad = f(y) - f(x) + \sum_{j=1}^{k+1}(f(x_j-) - f(x_{j-1}+)) \leq f(y) - f(x).\end{aligned}$$

Taking the supremum over all partitions of $[x,y]$ we have

$$f(x+)-f(x)+\sum_{\xi\in(x,y)}^{k}(f(x+)-f(x-))+f(y)-f(y-)\leq f(y)-f(x),$$

from which we deduce that

$$j_f(y)-j_f(x)=f(x+)-f(x)+\sum_{\xi\in(x,y)}^{k}(f(x+)-f(x-))+f(y)-f(y-)\leq f(y)-f(x).$$

This shows that $j_f(y)\geq j_f(x)$ and that $f(y)-j_f(y)\geq f(x)-j_f(x)$, showing that j_f and $f-j_f$ are monotonically increasing.

Now note that, as we saw in the proof of Proposition 3.3.18,

$$\begin{aligned} j_f(x+)-j_f(x)&=r_{f,x}, && x\in[a,b),\\ j_f(x)-j_f(x-)&=l_{f,x}, && x\in(a,b]. \end{aligned}$$

We also have $j_f(a)=0$. Thus, for $x\in[a,b)$, we have

$$(f(x)-j_f(x))-(f(x-)-j_f(x-))=f(x)-f(x-)-l_{f,x}=0$$

and, for $x\in(a,b]$, we have

$$(f(x+)-j_f(x+))-(f(x)-j_f(x))=f(x+)-f(x)-r_{f,x}=0.$$

Thus $f-j_f$ is continuous, as claimed. ■

This gives the following corollary which follows more or less directly from Theorem 3.3.3(ii).

3.3.23 Corollary (Saltus function of a function of bounded variation) *Let* $\mathrm{I}=[\mathrm{a},\mathrm{b}]$ *be a compact interval and let* $\mathrm{f}\colon \mathrm{I}\to\mathbb{R}$ *be of bounded variation. Define two families* $(\mathrm{r}_{\mathrm{f},\mathrm{x}})_{\mathrm{x}\in\mathrm{I}}$, *and* $(\mathrm{l}_{\mathrm{f},\mathrm{x}})_{\mathrm{x}\in\mathrm{I}}$ *of real numbers by*

$$\begin{aligned} &\mathrm{r}_{\mathrm{f},\mathrm{x}}=\mathrm{f}(\mathrm{x}+)-\mathrm{f}(\mathrm{x}), && \mathrm{x}\in[\mathrm{a},\mathrm{b}),\\ &\mathrm{l}_{\mathrm{f},\mathrm{a}}=0,\ \mathrm{l}_{\mathrm{f},\mathrm{x}}=\mathrm{f}(\mathrm{x})-\mathrm{f}(\mathrm{x}-), && \mathrm{x}\in(\mathrm{a},\mathrm{b}], \end{aligned}$$

and let $\mathrm{j}_\mathrm{f}\colon \mathrm{I}\to\mathbb{R}$ *be defined by*

$$\mathrm{j}_\mathrm{f}(\mathrm{x})=\sum_{\xi\in(-\infty,\mathrm{x})\cap\mathrm{I}}\mathrm{r}_{\mathrm{f},\xi}+\sum_{\xi\in(-\infty,\mathrm{x}]\cap\mathrm{I}}\mathrm{l}_{\mathrm{f},\xi}.$$

Then j_f *is a function of bounded variation, and the function* $\mathrm{f}-\mathrm{j}_\mathrm{f}$ *is a continuous function of bounded variation.*

Of course, the preceding two results carry over, with some notational complications at endpoints, to functions of locally bounded variation defined on general intervals.

Note that Examples 3.2.27 and 3.2.28 illustrate some of the features of saltus functions and functions of locally bounded variation. Indeed, the Cantor function of Example 3.2.27 is a function of locally bounded variation for which the associated saltus function is zero, while the function of Example 3.2.28 is "all" saltus function. Perhaps it is also useful to give a more mundane example to illustrate the decomposition of a function of locally bounded variation into its saltus and continuous part.

3.3.24 Example (Saltus function of a function of locally bounded variation) Let $I = [0,1]$ and consider three functions $f_1, f_2, f_3 \colon I \to \mathbb{R}$ defined by

$$f_1(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}], \\ -1, & x \in (\frac{1}{2}, 1], \end{cases}$$

$$f_2(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}], \\ 0, & x = \frac{1}{2}, \\ -1, & x \in (\frac{1}{2}, 1], \end{cases}$$

$$f_3(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}), \\ -1, & x \in [\frac{1}{2}, 1]. \end{cases}$$

In Example 3.3.5–3 we explicitly showed that f_1 is a function of locally bounded variation, and a similar argument shows that f_2 and f_3 are also functions of locally bounded variation. A direct application of the definition of Corollary 3.3.23 gives

$$j_{f_1}(x) = \begin{cases} 0, & x \in [0, \frac{1}{2}], \\ -2, & x \in (\frac{1}{2}, 1], \end{cases}$$

$$j_{f_2}(x) = \begin{cases} 0, & x \in [0, \frac{1}{2}), \\ -1, & x = \frac{1}{2}, \\ -2, & x \in (\frac{1}{2}, 1], \end{cases}$$

$$j_{f_3}(x) = \begin{cases} 0, & x \in [0, \frac{1}{2}), \\ -2, & x \in [\frac{1}{2}, 1]. \end{cases}$$

For $k \in \{1,2,3\}$ we have $f_k(x) = j_{f_k}(x) = 1$, $x \in [0,1]$. •

One might think that this is all that can be done as far as goes the decomposition of a function with locally bounded variation. However, this is not so. However, to further refine our present decomposition requires the notion of the integral as we consider it in Chapter 5. Thus we postpone a more detailed discussion of functions of locally bounded variation until ***missing stuff***.

Exercises

3.3.1 Show that if $I \subseteq \mathbb{R}$ is an interval and if $f \colon I \to \mathbb{R}$ is continuous then the following statements are equivalent:

1. f is injective;
2. f is either strictly monotonically increasing or strictly monotonically decreasing.

3.3.2 On the interval $I = [-1,1]$ consider the function $f \colon I \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{2}x + x^2 \sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

(a) Show that f is differentiable at $x = 0$ and has a positive derivative there.

(b) Show that for every $\epsilon \in \mathbb{R}_{>0}$ the restriction of f to $[-\epsilon, \epsilon]$ is neither monotonically decreasing (not surprisingly) nor monotonically increasing (surprisingly).

(c) Why is this not in contradiction with Proposition 3.2.23?

3.3.3 Give an example of an interval I and a function $f\colon I \to \mathbb{R}$ that is continuous, strictly monotonically increasing, but not differentiable.

3.3.4 Prove the assertions of Remark 3.3.7.

3.3.5 Let I be an interval and suppose that $I = I_1 \cup I_2$ where $I_1 \cap I_2 = \{x_0\}$ for some $x_0 \in \mathbb{R}$. If $f\colon I \to \mathbb{F}$ then

$$\mathrm{V}(f)(x) = \begin{cases} \mathrm{V}(f|I_1)(x), & x \in I_1, \\ \mathrm{V}(f|I_2)(x) + \mathrm{V}(f|I_1)(x_0), & x \in I_2 \end{cases}$$

if I_1 is finite,

$$\mathrm{V}(f)(x) = \begin{cases} \mathrm{V}(f|I_1)(x) - \mathrm{V}(f|I_2)(x_0), & x \in I_1, \\ \mathrm{V}(f|I_2)(x), & x \in I_2 \end{cases}$$

if I_1 is infinite and $x_0 < 0$, and

$$\mathrm{V}(f)(x) = \begin{cases} \mathrm{V}(f|I_1)(x), & x \in I_1, \\ \mathrm{V}(f|I_2)(x) + \mathrm{V}(f|I_1)(x_0), & x \in I_2 \end{cases}$$

if I_1 is infinite and $x_0 \geq 0$.

Section 3.4

The Riemann integral

Opposite to the derivative, in a sense made precise by Theorem 3.4.30, is the notion of integration. In this section we describe a "simple" theory of integration, called Riemann integration,[10] that typically works insofar as computations go. In Chapter 5 we shall see that the Riemann integration suffers from a defect somewhat like the defect possessed by rational numbers. That is to say, just like there are sequences of rational numbers that seem like they should converge (i.e., are Cauchy) but do not, there are sequences of functions possessing a Riemann integral which do not converge to a function possessing a Riemann integral (see Example 5.1.11). This has some deleterious consequences for developing a general theory based on the Riemann integral, and the most widely used fix for this is the Lebesgue integral of Chapter 5. However, for now let us stick to the more pedestrian, and more easily understood, Riemann integral.

As we did with differentiation, we suppose that the reader has had the sort of calculus course where they learn to compute integrals of common functions. Indeed, while we do not emphasise the art of computing integrals, we do not intend this to mean that this art should be ignored. The reader should know the basic integrals and the basic tricks and techniques for computing them. ***missing stuff***

Do I need to read this section? The best way to think of this section is as a setup for the general developments of Chapter 5. Indeed, we begin Chapter 5 with essentially a deconstruction of what we do in this section. For this reason, this chapter should be seen as preparatory to Chapter 5, and so can be skipped until one wants to learn Lebesgue integration in a serious way. At that time, a reader may wish to be prepared by understanding the slightly simpler Riemann integral. •

3.4.1 Step functions

Our discussion begins by our considering intervals that are compact. In Section 3.4.4 we consider the case of noncompact intervals.

In a theme that will be repeated when we consider the Lebesgue integral in Chapter 5, we first introduce a simple class of functions whose integral is "obvious." These functions are then used to approximate a more general class of functions which are those that are considered "integrable." For the Riemann integral, the simple class of functions are defined as being constant on the intervals forming a partition. We recall from Definition 2.5.7 the notion of a partition and from the

[10] After Georg Friedrich Bernhard Riemann, 1826–1866. Riemann made important and long lasting contributions to real analysis, geometry, complex function theory, and number theory, to name a few areas. The presently unsolved Riemann Hypothesis is one of the outstanding problems in modern mathematics.

discussion surrounding the definition the notion of the endpoints associated with a partition.

3.4.1 Definition (Step function) Let $I = [a, b]$ be a compact interval. A function $f\colon I \to \mathbb{R}$ is a ***step function*** if there exists a partition $P = (I_1, \dots, I_k)$ of I such that

(i) $f|\operatorname{int}(I_j)$ is a constant function for each $j \in \{1, \dots, k\}$,

(ii) $f(a+) = f(a)$ and $f(b-) = f(b)$, and

(iii) for each $x \in \mathrm{EP}(P) \setminus \{a, b\}$, either $f(x-) = f(x)$ or $f(x+) = f(x)$. •

In Figure 3.10 we depict a typical step function. Note that at discontinuities we allow the function to be continuous from either the right or the left. In the development we undertake, it does not really matter which it is.

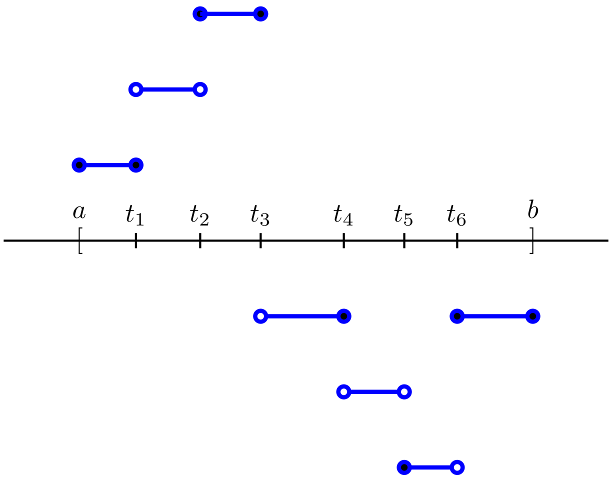


Figure 3.10 A step function

The idea of the integral of a function is that it measures the "area" below the graph of a function. If the value of the function is negative, then the area is taken to be negative. For step functions, this idea of the area under the graph is clear, so we simply define this to be the integral of the function.

3.4.2 Definition (Riemann integral of a step function) Let $I = [a, b]$ and let $f\colon I \to \mathbb{R}$ be a step function defined using the partition $P = (I_1, \dots, I_k)$ with endpoints $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$. Suppose that the value of f on $\operatorname{int}(I_j)$ is c_j for $j \in \{1, \dots, k\}$. The ***Riemann integral*** of f is

$$A(f) = \sum_{j=1}^{k} c_j(x_j - x_{j-1}). \qquad \bullet$$

The notation $A(f)$ is intended to suggest "area."

3.4.2 The Riemann integral on compact intervals

Next we define the Riemann integral of a function that is not necessarily a step function. We do this by approximating a function by step functions.

3.4.3 Definition (Lower and upper step functions) Let $I = [a, b]$ be a compact interval, let $f: I \to \mathbb{R}$ be a bounded function, and let $P = (I_1, \dots, I_k)$ be a partition of I.

(i) The ***lower step function*** associated to f and P is the function $s_-(f, P): I \to \mathbb{R}$ defined according to the following:

(a) if $x \in I$ lies in the interior of an interval I_j, $j \in \{1, \dots, k\}$, then $s_-(f, P)(x) = \inf\{f(x) \mid x \in \mathrm{cl}(I_j)\}$;

(b) $s_-(f, P)(a) = s_-(f, P)(a+)$ and $s_-(f, P)(b) = s_-(f, P)(b-)$;

(c) for $x \in \mathrm{EP}(P) \setminus \{a, b\}$, $s_-(f, P)(x) = s_-(f, P)(x+)$.

(ii) The ***upper step function*** associated to f and P is the function $s_+(f, P): I \to \mathbb{R}$ defined according to the following:

(a) if $x \in I$ lies in the interior of an interval I_j, $j \in \{1, \dots, k\}$, then $s_+(f, P)(x) = \sup\{f(x) \mid x \in \mathrm{cl}(I_j)\}$;

(b) $s_+(f, P)(a) = s_+(f, P)(a+)$ and $s_+(f, P)(b) = s_+(f, P)(b-)$;

(c) for $x \in \mathrm{EP}(P) \setminus \{a, b\}$, $s_+(f, P)(x) = s_+(f, P)(x+)$. •

Note that both the lower and upper step functions are well-defined since f is bounded. Note also that at the middle endpoints for the partition, we ask that the lower and upper step functions be continuous from the right. This is an arbitrary choice. Finally, note that for each $x \in [a, b]$ we have

$$s_-(f, P)(x) \leq f(x) \leq s_+(f, P)(x).$$

That is to say, for any bounded function f, we have defined two step functions, one bounding f from below and one bounding f from above.

Next we associate to the lower and upper step functions their integrals, which we hope to use to define the integral of the function f.

3.4.4 Definition (Lower and upper Riemann sums) Let $I = [a, b]$ be a compact interval, let $f: I \to \mathbb{R}$ be a bounded function, and let $P = (I_1, \dots, I_k)$ be a partition of I.

(i) The ***lower Riemann sum*** associated to f and P is $A_-(f, P) = A(s_-(f, P))$.

(ii) The ***upper Riemann sum*** associated to f and P is $A_+(f, P) = A(s_+(f, P))$. •

Now we define the best approximations of the integral of f using the lower and upper Riemann sums.

3.4.5 Definition (Lower and upper Riemann integral) Let $I = [a, b]$ be a compact interval and let $f: I \to \mathbb{R}$ be a bounded function.

(i) The ***lower Riemann integral*** of f is

$$I_-(f) = \sup\{A_-(f, P) \mid P \in \mathrm{Part}(I)\}.$$

(ii) The ***upper Riemann integral*** of f is

$$I_+(f) = \inf\{A_+(f,P) \mid P \in \mathrm{Part}(I)\}.$$ •

Note that since f is bounded, it follows that the sets

$$\{A_-(f,P) \mid P \in \mathrm{Part}(I)\}, \quad \{A_+(f,P) \mid P \in \mathrm{Part}(I)\}$$

are bounded (why?). Therefore, the lower and upper Riemann integral always exist. So far, then, we have made a some constructions that apply to *any* bounded function. That is to say, for any bounded function, it is possible to define the lower and upper Riemann integral. What is not clear is that these two things should be equal. In fact, they are *not* generally equal, which leads to the following definition.

3.4.6 Definition (Riemann integrable function on a compact interval) A bounded function $f\colon [a,b] \to \mathbb{R}$ on a compact interval is ***Riemann integrable*** if $I_-(f) = I_+(f)$. We denote

$$\int_a^b f(x)\,\mathrm{d}x = I_-(f) = I_+(f),$$

which is the ***Riemann integral*** of f. The function f is called the ***integrand***. •

3.4.7 Notation (Swapping limits of integration) In the expression $\int_a^b f(x)\,\mathrm{d}x$, "$a$" is the ***lower limit of integration*** and "b" is the ***upper limit of integration***. We have tacitly assumed that $a < b$ in our constructions to this point. However, we can consider the case where $b < a$ by adopting the convention that

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x.$$ •

Let us provide an example which illustrates that, in principle, it is possible to use the definition of the Riemann integral to perform computations, even though this is normally tedious. A more common method for computing integrals is to use the Fundamental Theorem of Calculus to "reverse engineer" the process.

3.4.8 Example (Computing a Riemann integral) Let $I = [0,1]$ and define $f\colon I \to \mathbb{R}$ by $f(x) = x$. Let $P = (I_1, \dots, I_k)$ be a partition with $s_-(f,P)$ and $s_+(f,P)$ the associated lower and upper step functions, respectively. Let $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$ be the endpoints of the intervals of the partition. One can then see that, for $j \in \{1, \dots, k\}$, $s_-(f,P)|\mathrm{int}(I_j) = x_{j-1}$ and $s_+(f,P)|\mathrm{int}(I_j) = x_j$. Therefore,

$$A_-(f,P) = \sum_{j=1}^k x_{j-1}(x_j - x_{j-1}), \quad A_+(f,P) = \sum_{j=1}^k x_j(x_j - x_{j-1}).$$

We claim that $I_-(f) \geq \frac{1}{2}$ and that $I_+(f) \leq \frac{1}{2}$, and note that, once we prove this, it follows that f is Riemann integrable and that $I_-(f) = I_+(f) = \frac{1}{2}$ (why?).

For $k \in \mathbb{Z}_{>0}$ consider the partition P_k with endpoints $\mathrm{EP}(P_k) = \{\frac{j}{k} \mid j \in \{0, 1, \dots, k\}\}$. Then, using the formula $\sum_{j=1}^{l} j = \frac{1}{2}l(l+1)$, we compute

$$A_-(f, P_k) = \sum_{j=1}^{k} \frac{j-1}{k^2} = \frac{k(k-1)}{2k^2}, \qquad A_+(f, P_k) = \sum_{j=1}^{k} \frac{j}{k^2} = \frac{k(k+1)}{2k^2}.$$

Therefore,

$$\lim_{k\to\infty} A_-(f, P_k) = \tfrac{1}{2}, \qquad \lim_{k\to\infty} A_+(f, P_k) = \tfrac{1}{2}.$$

This shows that $I_-(f) \geq \frac{1}{2}$ and that $I_+(f) \leq \frac{1}{2}$, as desired. •

3.4.3 Characterisations of Riemann integrable functions on compact intervals

In this section we provide some insightful characterisations of the notion of Riemann integrability. First we provide four equivalent characterisations of the Riemann integral. Each of these captures, in a slightly different manner, the notion of the Riemann integral as a limit. It will be convenient to introduce the language that a ***selection*** from a partition $P = (I_1, \dots, I_k)$ is a family $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)$ of points such that $\xi_j \in \mathrm{cl}(I_j)$, $j \in \{1, \dots, k\}$.

3.4.9 Theorem (Riemann, Darboux,[11] and Cauchy characterisations of Riemann integrable functions) *For a compact interval* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *and a bounded function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$, *the following statements are equivalent:*

(i) f *is Riemann integrable;*

(ii) *for every* $\epsilon \in \mathbb{R}_{>0}$, *there exists a partition* P *such that* $\mathrm{A}_+(\mathrm{f}, \mathrm{P}) - \mathrm{A}_-(\mathrm{f}, \mathrm{P}) < \epsilon$ *(**Riemann's condition**);*

(iii) *there exists* $\mathrm{I}(\mathrm{f}) \in \mathbb{R}$ *such that, for every* $\epsilon \in \mathbb{R}_{>0}$ *there exists* $\delta \in \mathbb{R}_{>0}$ *such that, if* $\mathrm{P} = (\mathrm{I}_1, \dots, \mathrm{I}_\mathrm{k})$ *is a partition for which* $|\mathrm{P}| < \delta$ *and if* $(\xi_1, \dots, \xi_\mathrm{k})$ *is a selection from* P, *then*

$$\Bigl|\sum_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{f}(\xi_\mathrm{j})(\mathrm{x}_\mathrm{j} - \mathrm{x}_{\mathrm{j}-1}) - \mathrm{I}(\mathrm{f})\Bigr| < \epsilon,$$

where $\mathrm{EP}(\mathrm{P}) = (\mathrm{x}_0, \mathrm{x}_1, \dots, \mathrm{x}_\mathrm{k})$ *(**Darboux' condition**);*

(iv) *for each* $\epsilon \in \mathbb{R}_{>0}$ *there exists* $\delta \in \mathbb{R}_{>0}$ *such that, for any partitions* $\mathrm{P} = (\mathrm{I}_1, \dots, \mathrm{I}_\mathrm{k})$ *and* $\mathrm{P}' = (\mathrm{I}'_1, \dots, \mathrm{I}'_{\mathrm{k}'})$ *with* $|\mathrm{P}|, |\mathrm{P}'| < \delta$ *and for any selections* $(\xi_1, \dots, \xi_\mathrm{k})$ *and* $(\xi'_1, \dots, \xi'_{\mathrm{k}'})$ *from* P *and* P′, *respectively, we have*

$$\Bigl|\sum_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{f}(\xi_\mathrm{j})(\mathrm{x}_\mathrm{j} - \mathrm{x}_{\mathrm{j}-1}) - \sum_{\mathrm{j}=1}^{\mathrm{k}'} \mathrm{f}(\xi'_\mathrm{j})(\mathrm{x}'_\mathrm{j} - \mathrm{x}'_{\mathrm{j}-1})\Bigr| < \epsilon,$$

where $\mathrm{EP}(\mathrm{P}) = (\mathrm{x}_0, \mathrm{x}_1, \dots, \mathrm{x}_\mathrm{k})$ *and* $\mathrm{EP}(\mathrm{P}') = (\mathrm{x}'_0, \mathrm{x}'_1, \dots, \mathrm{x}'_{\mathrm{k}'})$ *(**Cauchy's condition**).*

Proof First let us prove a simple lemma about lower and upper Riemann sums and refinements of partitions.

[11] Jean Gaston Darboux (1842–1917) was a French mathematician. His made important contributions to analysis and differential geometry.

1 Lemma *Let* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$*, let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be bounded, and let* P_1 *and* P_2 *be partitions of* I *with* P_2 *a refinement of* P_1*. Then*

$$\mathrm{A}_-(\mathrm{f}, \mathrm{P}_2) \geq \mathrm{A}_-(\mathrm{f}, \mathrm{P}_1), \quad \mathrm{A}_+(\mathrm{f}, \mathrm{P}_2) \leq \mathrm{A}_+(\mathrm{f}, \mathrm{P}_1).$$

Proof Let $x_1, x_2 \in \mathrm{EP}(P_1)$ and denote by $y_1, \dots, y_l$ the elements of $\mathrm{EP}(P_2)$ that satisfy

$$x_1 \leq y_1 < \cdots < y_l \leq x_2.$$

Then

$$\begin{aligned}\sum_{j=1}^{l}(y_j - y_{j-1}) \inf\{f(y) \mid y \in [y_j, y_{j-1}]\} &\geq \sum_{j=1}^{l}(y_j - y_{j-1}) \inf\{f(x) \mid x \in [x_1, x_2]\} \\ &= (x_2 - x_1) \inf\{f(x) \mid x \in [x_1, x_2]\}.\end{aligned}$$

Now summing over all consecutive pairs of endpoints for P_1 gives $A_-(f, P_2) \geq A_-(f, P_1)$. A similar argument gives $A_+(f, P_2) \leq A_+(f, P_1)$. ▼

The following trivial lemma will also be useful.

2 Lemma $\mathrm{I}_-(\mathrm{f}) \leq \mathrm{I}_+(\mathrm{f})$.

Proof Since, for any two partitions P_1 and P_2, we have

$$s_-(f, P_1) \leq f(x) \leq s_+(f, P_2),$$

it follows that

$$\sup\{A_-(f, P) \mid P \in \mathrm{Part}(I)\} \leq \inf\{A_+(f, P) \mid P \in \mathrm{Part}(I)\},$$

which is the result. ▼

(i) $\Longrightarrow$ (ii) Suppose that f is Riemann integrable and let $\epsilon \in \mathbb{R}_{>0}$. Then there exists partitions P_- and P_+ such that

$$A_-(f, P_-) > I_-(f) - \tfrac{\epsilon}{2}, \quad A_+(f, P_+) < I_+(f) + \tfrac{\epsilon}{2}.$$

Now let P be a partition that is a refinement of both P_1 and P_2 (obtained, for example, by asking that $\mathrm{EP}(P) = \mathrm{EP}(P_1) \cup \mathrm{EP}(P_2)$). By Lemma 1 it follows that

$$A_+(f, P) - A_-(f, P) \leq A_+(f, P_+) - A_-(f, P_-) < I_+(f) + \tfrac{\epsilon}{2} - I_-(f) + \tfrac{\epsilon}{2} = \epsilon.$$

(ii) $\Longrightarrow$ (i) Now suppose that $\epsilon \in \mathbb{R}_{>0}$ and let P be a partition such that $A_+(f, P) - A_-(f, P) < \epsilon$. Since we additionally have $I_-(f) \leq I_+(f)$ by Lemma 2, it follows that

$$A_-(f, P) \leq I_-(f) \leq I_+(f) \leq A_+(f, P),$$

from which we deduce that

$$0 \leq I_+(f) - I_-(f) < \epsilon.$$

Since ϵ is arbitrary, we conclude that $I_-(f) = I_+(f)$, as desired.

(i) $\Longrightarrow$ (iii) We first prove a lemma about partitions of compact intervals.

3 Lemma *If* $\mathrm{P} = (\mathrm{I}_1, \dots, \mathrm{I}_\mathrm{k})$ *is a partition of* $[\mathrm{a}, \mathrm{b}]$ *and if* $\epsilon \in \mathbb{R}_{>0}$*, then there exists* $\delta \in \mathbb{R}_{>0}$ *such that, if* $\mathrm{P}' = (\mathrm{I}'_1, \dots, \mathrm{I}'_{\mathrm{k}'})$ *is a partition with* $|\mathrm{P}'| < \delta$ *and if*

$$\{\mathrm{j}'_1, \dots, \mathrm{j}'_\mathrm{r}\} = \{\mathrm{j}' \in \{1, \dots, \mathrm{k}'\} \mid \mathrm{cl}(\mathrm{I}'_{\mathrm{j}'}) \not\subset \mathrm{cl}(\mathrm{I}_\mathrm{j}) \textit{ for any } \mathrm{j} \in \{1, \dots, \mathrm{k}\}\},$$

then

$$\sum_{\mathrm{l}=1}^{\mathrm{r}} |\mathrm{x}_{\mathrm{j}'_\mathrm{l}} - \mathrm{x}_{\mathrm{j}'_\mathrm{l}-1}| < \epsilon,$$

where $\mathrm{EP}(\mathrm{P}') = (\mathrm{x}_0, \mathrm{x}_1, \dots, \mathrm{x}_{\mathrm{k}'})$.

Proof Let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \frac{\epsilon}{k+1}$. Let $P' = (I'_1, \dots, I'_{k'})$ be a partition with endpoints $(x_0, x_1, \dots, x_{k'})$ and satisfying $|P'| < \delta$. Define

$$K_1 = \{j' \in \{1, \dots, k'\} \mid \mathrm{cl}(I'_{j'}) \not\subset \mathrm{cl}(I_j) \text{ for any } j \in \{1, \dots, k\}\}.$$

If $j' \in K_1$ then $I'_{j'}$ is not contained in any interval of P and so $I'_{j'}$ must contain at least one endpoint from P. Since P has $k+1$ endpoints we obtain $\mathrm{card}(K_1) \leq k+1$. Since the intervals $I'_{j'}$, $j' \in K_1$, have length at most δ we have

$$\sum_{j' \in K_1} (x_{j'} - x_{j'-1}) \leq (k+1)\delta \leq \epsilon,$$

as desired. ▼

Now let $\epsilon \in \mathbb{R}_{>0}$ and define $M = \sup\{|f(x)| \mid x \in I\}$. Denote by $I(f)$ the Riemann integral of f. Choose partitions P_- and P_+ such that

$$I(f) - A_-(f, P_-) < \tfrac{\epsilon}{2}, \quad A_+(f, P_+) - I(f) < \tfrac{\epsilon}{2}.$$

If $P = (I_1, \dots, I_k)$ is chosen such that $\mathrm{EP}(P) = \mathrm{EP}(P_-) \cup \mathrm{EP}(P_+)$, then

$$I(f) - A_-(f, P) < \tfrac{\epsilon}{2}, \quad A_+(f, P) - I(f) < \tfrac{\epsilon}{2}.$$

By Lemma 3 choose $\delta \in \mathbb{R}_{>0}$ such that if P' is any partition for which $|P'| < \delta$ then the sum of the lengths of the intervals of P' not contained in some interval of P does not exceed $\frac{\epsilon}{2M}$. Let $P' = (I'_1, \dots, I'_{k'})$ be a partition with endpoints $(x_0, x_1, \dots, x_{k'})$ and satisfying $|P'| < \delta$. Denote

$$K_1 = \{j' \in \{1, \dots, k'\} \mid I'_{j'} \not\subset I_j \text{ for some } j \in \{1, \dots, k\}\}$$

and $K_2 = \{1, \dots, k'\} \setminus K_1$. Let $(\xi_1, \dots, \xi_{k'})$ be a selection of P'. Then we compute

$$\begin{aligned}\sum_{j=1}^{k'} f(\xi_j)(x_j - x_{j-1}) &= \sum_{j \in K_1} f(\xi_j)(x_j - x_{j-1}) + \sum_{j \in K_2} f(\xi_j)(x_j - x_{j-1}) \\ &\leq A_+(f, P) + M\frac{\epsilon}{2M} < I(f) + \epsilon.\end{aligned}$$

In like manner we show that

$$\sum_{j=1}^{k'} f(\xi_j)(x_j - x_{j-1}) > I(f) - \epsilon.$$

This gives

$$\Big|\sum_{j=1}^{k'} f(\xi_j)(x_j - x_{j-1}) - I(f)\Big| < \epsilon,$$

as desired.

(iii) $\implies$ (ii) Let $\epsilon \in \mathbb{R}_{>0}$ and let $P = (I_1, \dots, I_k)$ be a partition for which

$$\Big|\sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1}) - I(f)\Big| < \frac{\epsilon}{4}$$

for every selection $(\xi_1, \dots, \xi_k)$ from P. Now particularly choose a selection such that

$$|f(\xi_j) - \sup\{f(x) \mid x \in \mathrm{cl}(I_j)\}| < \frac{\epsilon}{4k(x_j - x_{j-1})}.$$

Then

$$\begin{aligned}|A_+(f,P) - I(f)| &\le \Big|A_+(f,P) - \sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1})\Big| + \Big|\sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1}) - I(f)\Big| \\ &< \sum_{j=1}^{k} \frac{\epsilon}{4k(x_j - x_{j-1})}(x_j - x_{j-1}) + \frac{\epsilon}{4} < \frac{\epsilon}{2}.\end{aligned}$$

In like manner one shows that $|A_-(f,P) - I(f)| < \frac{\epsilon}{2}$. Therefore,

$$|A_+(f,P) - A_-(f,P)| \le |A_+(f,P) - I(f)| + |I(f) - A_-(f,P)| < \epsilon,$$

as desired.

(iii) $\implies$ (iv) Let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta \in \mathbb{R}_{>0}$ have the property that, whenever $P = (I_1, \dots_k)$ is a partition satisfying $|P| < \delta$ and $(\xi_1, \dots, \xi_k)$ is a selection from P, it holds that

$$\Big|\sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1}) - I(f)\Big| < \frac{\epsilon}{2}.$$

Now let $P = (I_1, \dots, I_k)$ and $P' = (I'_1, \dots, I'_{k'})$ be two partitions with $|P|, |P'| < \delta$, and let $(\xi_1, \dots, \xi_k)$ and $(\xi'_1, \dots, \xi'_{k'})$ selections from P and P', respectively. Then we have

$$\begin{aligned}&\Big|\sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1}) - \sum_{j=1}^{k'} f(\xi'_j)(x'_j - x'_{j-1})\Big| \\ &\qquad \le \Big|\sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1}) - I(f)\Big| + \Big|\sum_{j=1}^{k'} f(\xi'_j)(x'_j - x'_{j-1}) - I(f)\Big| < \epsilon,\end{aligned}$$

which gives this part of the result.

(iv) $\implies$ (iii) Let $(P_j = (I_{j,1}, \dots, I_{j,k_j}))_{j\in\mathbb{Z}_{>0}}$ be a sequence of partitions for which $\lim_{j\to\infty}|P_j| = 0$. Then, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that

$$\Big|\sum_{j=1}^{k_l} f(\xi_{l,j})(x_{l,j} - x_{l,j-1}) - \sum_{j=1}^{k_m} f(\xi_{m,j})(x_{m,j} - x_{m,j-1})\Big| < \epsilon,$$

for $l, m \geq N$, where $\boldsymbol{\xi}_j = (\xi_{j,1}, \dots, \xi_{j,k_j})$, is a selection from P_j, $j \in \mathbb{Z}_{>0}$, and where $\mathrm{EP}(P_j) = (x_{j,0}, x_{j,1}, \dots, x_{j,k_j})$, $j \in \mathbb{Z}_{>0}$. If we define

$$A(f, P_j, \boldsymbol{\xi}_j) = \sum_{r=1}^{k_j} f(\xi_r)(x_{j,r} - x_{j,r-1}),$$

then the sequence $(A(f, P_j, \boldsymbol{\xi}_j))_{j \in \mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{R}$ for any choices of points $\boldsymbol{\xi}_j$, $j \in \mathbb{Z}_{>0}$. Denote the resulting limit of this sequence by $I(f)$. We claim that $I(f)$ is the Riemann integral of f. To see this, let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta \in \mathbb{R}_{>0}$ be such that

$$\Bigl|\sum_{j=1}^{k} f(\xi_j)(x_j - x_{j-1}) - \sum_{j=1}^{k'} f(\xi'_j)(x'_j - x'_{j-1})\Bigr| < \frac{\epsilon}{2}$$

for any two partitions P and P' satisfying $|P|, |P'| < \delta$ and for any selections $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ from P and P', respectively. Now let $N \in \mathbb{Z}_{>0}$ satisfy $|P_j| < \delta$ for every $j \geq N$. Then, if P is any partition with $|P| < \delta$ and if $\boldsymbol{\xi}$ is any selection from P, we have

$$|A(f, P, \boldsymbol{\xi}) - I(f)| \leq |A(f, P, \boldsymbol{\xi}) - A(f, P_N, \boldsymbol{\xi}_N)| + |A(f, P_N, \boldsymbol{\xi}_N) - I(f)| < \epsilon,$$

for any selection $\boldsymbol{\xi}_N$ of P_N. This shows that $I(f)$ is indeed the Riemann integral of f, and so gives this part of the theorem. ■

A consequence of the proof is that, of course, the quantity $I(f)$ in part (iii) of the theorem is nothing other than the Riemann integral of f.

Many of the functions one encounters in practice are, in fact, Riemann integrable. However, not all functions are Riemann integrable, as the following simple examples shows.

3.4.10 Example (A function that is not Riemann integrable) Let $I = [0, 1]$ and let $f : I \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \cap I \\ 0, & x \notin \mathbb{Q} \cap I. \end{cases}$$

Thus f takes the value 1 at all rational points, and is zero elsewhere. Now let $s_+, s_- : I \to \mathbb{R}$ be any step functions satisfying $s_-(x) \leq f(x) \leq s_+(x)$ for all $x \in I$. Since any nonempty subinterval of I contains infinitely many irrational numbers, it follows that $s_-(x) \leq 0$ for every $x \in I$. Since every nonempty subinterval of I contains infinitely many rational numbers, it follows that $s_+(x) \geq 1$ for every $x \in I$. Therefore, $A(s_+) - A(s_-) \geq 1$. It follows from Theorem 3.4.9 that f is not Riemann integrable. While this example may seem pointless and contrived, it will be used in Examples **????** and 5.1.11 to exhibit undesirable features of the Riemann integral. •

The following result provides an interesting characterisation of Riemann integrable functions, illustrating precisely the sorts of functions whose Riemann integrals may be computed.

3.4.11 Theorem (Riemann integrable functions are continuous almost everywhere, and vice versa) *For a compact interval* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$*, a bounded function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is Riemann integrable if and only if the set*

$$\mathrm{D}_{\mathrm{f}} = \{\mathrm{x} \in \mathrm{I} \mid \mathrm{f} \textit{ is discontinuous at } \mathrm{x}\}$$

has measure zero.

Proof Recall from Definition 3.1.10 the notion of the oscillation ω_f for a function f, and that $\omega_f(x) = 0$ if and only if f is continuous at x. For $k \in \mathbb{Z}_{>0}$ define

$$D_{f,k} = \left\{x \in I \;\middle|\; \omega_f(x) \geq \tfrac{1}{k}\right\}.$$

Then Proposition 3.1.11 implies that $D_f = \cup_{k \in \mathbb{Z}_{>0}} D_{f,k}$. By Exercise 2.5.9 we can assert that D_f has measure zero if and only if each of the sets $D_{f,k}$ has measure zero, $k \in \mathbb{Z}_{>0}$.

Now suppose that $D_{f,k}$ does not have measure zero for some $k \in \mathbb{Z}_{>0}$. Then there exists $\epsilon \in \mathbb{R}_{>0}$ such that, if a family $((a_j, b_j))_{j \in \mathbb{Z}_{>0}}$ of open intervals has the property that

$$D_{f,k} \subseteq \bigcup_{j \in \mathbb{Z}_{>0}} (a_j, b_j),$$

then

$$\sum_{j=1}^{\infty} |b_j - a_j| \geq \epsilon.$$

Now let P be a partition of I and denote $\mathrm{EP}(P) = (x_0, x_1, \dots, x_m)$. Now let $\{j_1, \dots, j_l\} \subseteq \{1, \dots, m\}$ be those indices for which $j_r \in \{j_1, \dots, j_l\}$ implies that $D_{f,k} \cap (x_{j_r - 1}, x_{j_r}) \neq \emptyset$. Note that it follows that the set $\bigcup_{r=1}^{l} (x_{j_r - 1}, x_{j_r})$ covers $D_{f,k}$ with the possible exception of a finite number of points. It then follows that one can enlarge the length of each of the intervals $(x_{j_r - 1}, x_{j_r})$, $r \in \{1, \dots, l\}$, by $\frac{\epsilon}{2l}$, and the resulting intervals will cover $D_{f,k}$. The enlarged intervals will have total length at least ϵ, which means that

$$\sum_{r=1}^{l} |x_{j_r} - x_{j_r - 1}| \geq \frac{\epsilon}{2}.$$

Moreover, for each $r \in \{1, \dots, l\}$,

$$\sup\{f(x) \mid x \in [x_{j_r - 1}, x_{j_r}]\} - \inf\{f(x) \mid x \in [x_{j_r - 1}, x_{j_r}]\} \geq \tfrac{1}{k}$$

since $D_{f,k} \cap (x_{j_r - 1}, x_{j_r}) \neq \emptyset$ and by definition of $D_{f,k}$ and ω_f. It now follows that

$$\begin{aligned} A_+(f, P) - A_-(f, P) = {} & \sum_{j=1}^{m} (x_j - x_{j-1}) \big(\sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\ & - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\}\big) \\ \geq {} & \sum_{r=1}^{l} (x_{j_r} - x_{j_r - 1}) \big(\sup\{f(x) \mid x \in [x_{j_r - 1}, x_{j_r}]\} \\ & - \inf\{f(x) \mid x \in [x_{j_r - 1}, x_{j_r}]\}\big) \\ \geq {} & \tfrac{\epsilon}{2k}. \end{aligned}$$

Since this must hold for every partition, it follows that f is not Riemann integrable.

Now suppose that D_f has measure zero. Since f is bounded, let $M = \sup\{|f(x)| \mid x \in I\}$. Let $\epsilon \in \mathbb{R}_{>0}$ and for brevity define $\epsilon' = \frac{\epsilon}{b-a+2}$. Choose a sequence $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ of open intervals such that

$$D_f \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} I_j, \quad \sum_{j=1}^{\infty} |b_j - a_j| < \tfrac{\epsilon'}{M}.$$

Define $\delta\colon I \to \mathbb{R}_{>0}$ such that the following properties hold:

1. if $x \notin D_f$ then $\delta(x)$ is taken such that, if $y \in I \cap \mathsf{B}(\delta(x), x)$, then $|f(y) - f(x)| < \frac{\epsilon'}{2}$;
2. if $x \in D_f$ then $\delta(x)$ is taken such that $\mathsf{B}(\delta(x), x) \subseteq I_j$ for some $j \in \mathbb{Z}_{>0}$.

Now, by Proposition 2.5.10, let $((c_1, I_1), \dots, (c_k, I_k))$ be a δ-fine tagged partition with $P = (I_1, \dots, I_k)$ the associated partition. Now partition the set $\{1, \dots, k\}$ into two sets K_1 and K_2 such that $j \in K_1$ if and only if $c_j \notin D_f$. Then we compute

$$\begin{aligned}
A_+(f,P) - A_-(f,P) &= \sum_{j=1}^{k} (x_j - x_{j-1})\big(\sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\
&\quad - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\}\big) \\
&= \sum_{j\in K_1} (x_j - x_{j-1})\big(\sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\
&\quad - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\}\big) \\
&\quad + \sum_{j\in K_2} (x_j - x_{j-1})\big(\sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\
&\quad - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\}\big) \\
&\le \sum_{j\in K_1} \epsilon'(x_j - x_{j-1}) + \sum_{j\in K_2} 2M(x_j - x_{j-1}) \\
&\le \epsilon'(b-a) + 2M \sum_{j=1}^{\infty} |b_j - a_j| \\
&< \epsilon'(b - a + 2) = \epsilon.
\end{aligned}$$

This part of the result now follows by Theorem 3.4.9. ■

The theorem indicates why the function of Example 3.4.10 is not Riemann integrable. Indeed, the function in that example is discontinuous at *all* points in $[0, 1]$ (why?). The theorem also has the following obvious corollary which illustrates why so many functions in practice are Riemann integrable.

3.4.12 Corollary (Continuous functions are Riemann integrable) *If* $\mathrm{f}\colon [\mathrm{a}, \mathrm{b}] \to \mathbb{R}$ *is continuous, then it is Riemann integrable.*

By virtue of Theorem 3.3.3, we also have the following result, giving another large class of Riemann integrable functions, distinct from those that are continuous.

3.4.13 Corollary (Functions of bounded variation are Riemann integrable) *If* $f\colon [a,b] \to \mathbb{R}$ *has bounded variation, then* f *is Riemann integrable.*

3.4.4 The Riemann integral on noncompact intervals

Up to this point in this section we have only considered the Riemann integral for bounded functions defined on compact intervals. In this section we extend the notion of the Riemann integral to allow its definition for unbounded functions and for general intervals. There are complications that arise in this situation that do not arise in the case of a compact interval in that one has two possible notions of what one might call a Riemann integrable function. In all cases, we use the existing definition of the Riemann integral for compact intervals as our basis, and allow the other cases as limits.

3.4.14 Definition (Positive Riemann integrable function on a general interval) Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}_{\geq 0}$ be a function whose restriction to every compact subinterval of I is Riemann integrable.

(i) If $I = [a,b]$ then the Riemann integral of f is as defined in the preceding section.

(ii) If $I = (a,b]$ then define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{r_a \downarrow a} \int_{r_a}^b f(x)\,\mathrm{d}x.$$

(iii) If $I = [a,b)$ then define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{r_b \uparrow b} \int_a^{r_b} f(x)\,\mathrm{d}x.$$

(iv) If $I = (a,b)$ then define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{r_a \downarrow a} \int_{r_a}^c f(x)\,\mathrm{d}x + \lim_{r_b \uparrow b} \int_c^{r_b} f(x)\,\mathrm{d}x$$

for some $c \in (a,b)$.

(v) If $I = (-\infty,b]$ then define

$$\int_{-\infty}^b f(x)\,\mathrm{d}x = \lim_{R \to \infty} \int_{-R}^b f(x)\,\mathrm{d}x.$$

(vi) If $I = (-\infty,b)$ then define

$$\int_{-\infty}^b f(x)\,\mathrm{d}x = \lim_{R \to \infty} \int_{-R}^c f(x)\,\mathrm{d}x + \lim_{r_b \uparrow b} \int_c^{r_b} f(x)\,\mathrm{d}x$$

for some $c \in (-\infty,b)$.

(vii) If $I = [a, \infty)$ then define

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{R\to\infty} \int_a^R f(x)\,\mathrm{d}x.$$

(viii) If $I = (a, \infty)$ then define

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{r_a \downarrow a} \int_{r_a}^c f(x)\,\mathrm{d}x + \lim_{R\to\infty} \int_c^R f(x)\,\mathrm{d}x$$

for some $c \in (a, \infty)$.

(ix) If $I = \mathbb{R}$ then define

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \lim_{R\to\infty} \int_{-R}^c f(x)\,\mathrm{d}x + \lim_{R\to\infty} \int_c^R f(x)\,\mathrm{d}x$$

for some $c \in \mathbb{R}$.

If, for a given I and f, the appropriate of the above limits exists, then f is ***Riemann integrable*** on I, and the ***Riemann integral*** is the value of the limit. Let us denote by

$$\int_I f(x)\,\mathrm{d}x$$

the Riemann integral. •

One can easily show that where, in the above definitions, one must make a choice of c, the definition is independent of this choice (cf. Proposition 3.4.26).

The above definition is intended for functions taking nonnegative values. For more general functions we have the following definition.

3.4.15 Definition (Riemann integrable function on a general interval) Let $I \subseteq \mathbb{R}$ be an interval and let $f: I \to \mathbb{R}$ be a function whose restriction to any compact subinterval of I is Riemann integrable. Define $f_+, f_-: I \to \mathbb{R}_{\geq 0}$ by

$$f_+(x) = \max\{0, f(x)\}, \quad f_-(x) = -\min\{0, f(x)\}$$

so that $f = f_+ - f_-$. The function f is ***Riemann integrable*** if both f_+ and f_- are Riemann integrable, and the ***Riemann integral*** of f is

$$\int_I f(x)\,\mathrm{d}x = \int_I f_+(x)\,\mathrm{d}x - \int_I f_-(x)\,\mathrm{d}x. \quad \bullet$$

At this point, if I is compact, we have potentially competing definitions for the Riemann integral of a bounded function $I: f \to \mathbb{R}$. One definition is the direct one of Definition 3.4.6. The other definition involves computing the Riemann integral, as per Definition 3.4.6, of the positive and negative parts of f, and then take the difference of these. Let us resolve the equivalence of these two notions.

3.4.16 Proposition (Consistency of definition of Riemann integral on compact intervals) *Let* $\mathrm{I} = [\mathrm{a},\mathrm{b}]$*, let* $\mathrm{f}\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}$*, and let* $\mathrm{f}_+, \mathrm{f}_-\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}_{\geq 0}$ *be the positive and negative parts of* f*. Then the following two statements are equivalent:*

(i) f *is integrable as per Definition 3.4.6 with Riemann integral* $\mathrm{I}(\mathrm{f})$*;*

(ii) f_+ *and* f_- *are Riemann integrable as per Definition 3.4.6 with Riemann integrals* $\mathrm{I}(\mathrm{f}_+)$ *and* $\mathrm{I}(\mathrm{f}_-)$*.*

Moreover, if one, and therefore both, of parts (i) and (ii) hold, then $\mathrm{I}(\mathrm{f}) = \mathrm{I}(\mathrm{f}_+) - \mathrm{I}(\mathrm{f}_-)$*.*

Proof We shall refer ahead to the results of Section 3.4.5.

(i) $\Longrightarrow$ (ii) Define continuous functions $g_+, g_-\colon \mathbb{R} \to \mathbb{R}$ by

$$g_+(x) = \max\{0, x\}, \quad g_-(x) = -\min\{0, x\}$$

so that $f_+ = g_+ \circ f$ and $f_- = g_- \circ f$. By Proposition 3.4.23 (noting that the proof of that result is valid for the Riemann integral as per Definition 3.4.6) it follows that f_+ and f_- are Riemann integrable as per Definition 3.4.6.

(ii) $\Longrightarrow$ (i) Note that $f = f_+ - f_-$. Also note that the proof of Proposition 3.4.22 is valid for the Riemann integral as per Definition 3.4.6. Therefore, f is Riemann integrable as per Definition 3.4.6.

Now we show that $I(f) = I(f_+) - I(f_-)$. This, however, follows immediately from Proposition 3.4.22. ■

It is not uncommon to see the general integral as we have defined it called the ***improper Riemann integral***.

The preceding definitions may appear at first to be excessively complicated. The following examples illustrate the rationale behind the care taken in the definitions.

3.4.17 Examples (Riemann integral on a general interval)

1. Let $I = (0, 1]$ and let $f(x) = x^{-1}$. Then, if $r_a \in (0, 1)$, we compute the proper Riemann integral
$$\int_{r_a}^{1} f(x)\,\mathrm{d}x = -\log r_a,$$
where log is the natural logarithm. Since $\lim_{r_a \downarrow} \log r_a = -\infty$ this function is not Riemann integrable on $(0, 1]$.

2. Let $I = (0, 1]$ and let $f(x) = x^{-1/2}$. Then, if $r_a \in (0, 1)$, we compute the proper Riemann integral
$$\int_{r_a}^{1} f(x)\,\mathrm{d}x = 2 - 2\sqrt{r_a}.$$
In this case the function is Riemann integrable on $(0, 1]$ and the value of the Riemann integral is 2.

3. Let $I = \mathbb{R}$ and define $f(x) = (1 + x^2)^{-1}$. In this case we have
$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,\mathrm{d}x &= \lim_{R\to\infty} \int_{-R}^{0} \frac{1}{1+x^2}\,\mathrm{d}x + \lim_{R\to\infty} \int_{0}^{R} \frac{1}{1+x^2}\,\mathrm{d}x \\ &= \lim_{R\to\infty} \arctan R + \lim_{R\to\infty} \arctan R = \pi.\end{aligned}$$
Thus this function is Riemann integrable on $\mathbb{R}$ and has a Riemann integral of π.

4. The next example we consider is $I = \mathbb{R}$ and $f(x) = x(1+x^2)^{-1}$. In this case we compute
$$\begin{aligned}\int_{-\infty}^{\infty} \frac{x}{1+x^2}\,\mathrm{d}x &= \lim_{R\to\infty}\int_{-R}^{0} \frac{x}{1+x^2}\,\mathrm{d}x + \lim_{R\to\infty}\int_{0}^{R} \frac{x}{1+x^2}\,\mathrm{d}x \\ &= \lim_{R\to\infty}\frac{1}{2}\log(1+R^2) - \lim_{R\to\infty}\frac{1}{2}\log(1+R^2).\end{aligned}$$
Now, it is not permissible to say here that $\infty - \infty = 0$. Therefore, we are forced to conclude that f is not Riemann integrable on $\mathbb{R}$.
5. To make the preceding example a little more dramatic, and to more convincingly illustrate why we should not cancel the infinities, we take $I = \mathbb{R}$ and $f(x) = x^3$. Here we compute
$$\int_{-\infty}^{\infty} x^3\,\mathrm{d}x = \lim_{R\to\infty}\frac{1}{4}R^4 - \lim_{R\to\infty}\frac{1}{4}R^4.$$
In this case again we must conclude that f is not Riemann integrable on $\mathbb{R}$. Indeed, it seems unlikely that one *would* wish to conclude that such a function was Riemann integrable since it is so badly behaved as $|t| \to \infty$. However, if we reject this function as being Riemann integrable, we must also reject the function of Example 4, even though it is not as ill behaved as the function here. •

Note that the above constructions involved first separating a function into its positive and negative parts, and then integrating these separately. However, there is not *a priori* reason why we could not have defined the limits in Definition 3.4.14 directly, and not just for positive functions. One can do this in fact. However, as we shall see, the two ensuing constructions of the integral are not equivalent.

3.4.18 Definition (Conditionally Riemann integrable functions on a general interval) Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}$ be a function whose restriction to any compact subinterval of I is Riemann integrable. Then f is ***conditionally Riemann integrable*** if the limit in the appropriate of the nine cases of Definition 3.4.14 exists. This limit is called the ***conditional Riemann integral*** of f. If f is conditionally integrable we write
$$C\int_I f(x)\,\mathrm{d}x$$
as the conditional Riemann integral. •

missing stuff

Before we explain the differences between conditionally integrable and integrable functions via examples, let us provide the relationship between the two notions.

3.4.19 Proposition (Relationship between integrability and conditional integrability) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$, *then the following statements hold:*

(i) *if* f *is Riemann integrable then it is conditionally Riemann integrable;*

(ii) if I *is additionally compact then, if* f *is conditionally Riemann integrable it is Riemann integrable.*

Proof In the proof it is convenient to make use of the results from Section 3.4.5.

(i) Let f_+ and f_- be the positive and negative parts of f. Since f is Riemann integrable, then so are f_+ and f_- by Definition 3.4.15. Moreover, since Riemann integrability and conditional Riemann integrability are clearly equivalent for nonnegative functions, it follows that f_+ and f_- are conditionally Riemann integrable. Therefore, by Proposition 3.4.22, it follows that $f = f_+ - f_-$ is conditionally Riemann integrable.

(ii) This follows from Definition 3.4.15 and Proposition 3.4.16. ■

Let us show that conditional Riemann integrability and Riemann integrability are not equivalent.

3.4.20 Example (A conditionally Riemann integrable function that is not Riemann integrable) Let $I = [1,\infty)$ and define $f(x) = \frac{\sin x}{x}$. Let us first show that f is conditionally Riemann integrable. We have, using integration by parts (Proposition 3.4.28),

$$\begin{aligned}\int_1^\infty \frac{\sin x}{x}\,\mathrm{d}x &= \lim_{R\to\infty}\int_1^R \frac{\sin x}{x}\,\mathrm{d}x = \lim_{R\to\infty}\Big(-\frac{\cos x}{x}\Big|_1^R - \int_1^R \frac{\cos x}{x^2}\,\mathrm{d}x\Big)\\ &= \cos 1 - \lim_{R\to\infty}\int_1^R \frac{\cos x}{x^2}\,\mathrm{d}x.\end{aligned}$$

We claim that the last limit exists. Indeed,

$$\Big|\int_1^R \frac{\cos x}{x^2}\,\mathrm{d}x\Big| \le \int_1^R \frac{|\cos x|}{x^2}\,\mathrm{d}x \le \int_1^R \frac{1}{x^2}\,\mathrm{d}x = 1 - \frac{1}{R},$$

and the limit as $R \to \infty$ is then 1. This shows that the limit defining the conditional integral is indeed finite, and so f is conditionally Riemann integrable on $[1,\infty)$.

Now let us show that this function is not Riemann integrable. By Proposition 3.4.25, f is Riemann integrable if and only if $|f|$ is Riemann integrable. For $R > 0$ let $N_R \in \mathbb{Z}_{>0}$ satisfy $R \in [N_R\pi, (N_R+1)\pi]$. We then have

$$\begin{aligned}\int_1^R \Big|\frac{\sin x}{x}\Big|\,\mathrm{d}x &\ge \int_\pi^{N_R\pi} \Big|\frac{\sin x}{x}\Big|\,\mathrm{d}x\\ &\ge \sum_{j=1}^{N_R-1} \frac{1}{j\pi}\int_{j\pi}^{(j+1)\pi} |\sin x|\,\mathrm{d}x = \frac{2}{\pi}\sum_{j=1}^{N_R-1}\frac{1}{j}.\end{aligned}$$

By Example 2.4.2–2, the last sum diverges to ∞ as $N_R \to \infty$, and consequently the integral on the left diverges to ∞ as $R \to \infty$, giving the assertion. •

3.4.21 Remark ("Conditional Riemann integral" versus "Riemann integral") The previous example illustrates that one needs to exercise some care when talking about the Riemann integral. Adding to the possible confusion here is the fact that there is no established convention concerning what is intended when one says "Riemann integral." Many authors use "Riemann integrability" where we use "conditional Riemann integrability" and then use "absolute Riemann integrability" where we use "Riemann integrability." There is a good reason to do this.

1. One can think of integrals as being analogous to sums. When we talked about convergence of sums in Section 2.4 we used "convergence" to talk about that concept which, for the Riemann integral, is analogous to "conditional Riemann integrability" in our terminology. We used the expression "absolute convergence" for that concept which, for the Riemann integral, is analogous to "Riemann integrability" in our terminology. Thus the alternative terminology of "Riemann integrability" for "conditional Riemann integrability" and "absolute Riemann integrability" for "Riemann integrability" is more in alignment with the (more or less) standard terminology for sums.

However, there is also a good reason to use the terminology we use. However, the reasons here have to do with terminology attached to the Lebesgue integral that we discuss in Chapter 5. However, here is as good a place as any to discuss this.

2. For the Lebesgue integral, the most natural notion of integrability is analogous to the notion of "Riemann integrability" in our terminology. That is, the terminology "Lebesgue integrability" is a generalisation of "Riemann integrability." The notion of "conditional Riemann integrability" is not much discussed for the Lebesgue integral, so there is not so much an established terminology for this. However, if there were an established terminology it would be "conditional Lebesgue integrability."

In Table 3.1 we give a summary of the preceding discussion, noting that apart from overwriting some standard conventions, there is no optimal way to choose what language to use. Our motivation for the convention we use is that it is best that "Lebesgue integrability" should generalise "Riemann integrability." But it is necessary to understand what one is reading and what is intended in any case. •

Table 3.1 "Conditional" versus "absolute" terminology. In the top row we give our terminology, in the second row we give the alternative terminology for the Riemann integral, in the third row we give the analogous terminology for sums, and in the fourth row we give the terminology for the Lebesgue integral.

	Riemann integrable	conditionally Riemann integrable
Alternative	absolutely Riemann integrable	Riemann integrable
Sums	absolutely convergent	convergent
Lebesgue integral	Lebesgue integrable	conditionally Lebesgue integrable

3.4.5 The Riemann integral and operations on functions

In this section we consider the interaction of integration with the usual algebraic and other operations on functions. We will consider both Riemann integrability and conditional Riemann integrability. If we wish to make a statement that we intend to hold for both notions, we shall write "(conditionally) Riemann integrable" to connote this. We will also write

$$(C)\int_I f(x)\,dx$$

to denote either the Riemann integral or the conditional Riemann integral in cases where we wish for both to apply. The reader should also keep in mind that Riemann integrability and conditional Riemann integrability agree for compact intervals.

3.4.22 Proposition (Algebraic operations and the Riemann integral) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval, let* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *be (conditionally) Riemann integrable functions, and let* $\mathrm{c} \in \mathbb{R}$*. Then the following statements hold:*

(i) $\mathrm{f} + \mathrm{g}$ *is (conditionally) Riemann integrable and*

$$(C)\int_{\mathrm{I}}(\mathrm{f}+\mathrm{g})(\mathrm{x})\,\mathrm{dx} = (C)\int_{\mathrm{I}}\mathrm{f}(\mathrm{x})\,\mathrm{dx} + (C)\int_{\mathrm{I}}\mathrm{g}(\mathrm{x})\,\mathrm{dx};$$

(ii) cf *is (conditionally) Riemann integrable and*

$$(C)\int_{\mathrm{I}}(\mathrm{cf})(\mathrm{x})\,\mathrm{dx} = \mathrm{c}(C)\int_{\mathrm{I}}\mathrm{f}(\mathrm{x})\,\mathrm{dx};$$

(iii) if I *is additionally compact, then* fg *is Riemann integrable;*

(iv) if I *is additionally compact and if there exists* $\alpha \in \mathbb{R}_{>0}$ *such that* $\mathrm{g}(\mathrm{x}) \geq \alpha$ *for each* $\mathrm{x} \in \mathrm{I}$*, then* $\frac{\mathrm{f}}{\mathrm{g}}$ *is Riemann integrable.*

Proof (i) We first suppose that $I = [a, b]$ is a compact interval. Let $\epsilon \in \mathbb{R}_{>0}$ and by Theorem 3.4.9 we let P_f and P_g be partitions of $[a, b]$ such that

$$A_+(f, P_f) - A_-(f, P_f) < \tfrac{\epsilon}{2}, \quad A_+(g, P_g) - A_-(g, P_g) < \tfrac{\epsilon}{2},$$

and let P be a partition for which $(x_0, x_1, \dots, x_k) = \mathrm{EP}(P) = \mathrm{EP}(P_f) \cup \mathrm{EP}(P_g)$. Then, using Proposition 2.2.27,

$$\sup\{f(x) + g(x) \mid x \in [x_{j-1}, x_j]\} = \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} + \sup\{g(x) \mid x \in [x_{j-1}, x_j]\}$$

and

$$\inf\{f(x) + g(x) \mid x \in [x_{j-1}, x_j]\} = \inf\{f(x) \mid x \in [x_{j-1}, x_j]\} + \inf\{g(x) \mid x \in [x_{j-1}, x_j]\}$$

for each $j \in \{1, \dots, k\}$. Thus

$$A_+(f + g, P) - A_-(f + g, P) \leq A_+(f, P) + A_+(g, P) - A_-(f, P) - A_-(g, P) < \epsilon,$$

using Lemma 1 from the proof of Theorem 3.4.9. This shows that $f+g$ is Riemann integrable by Theorem 3.4.9.

Now let P_f and P_g be any two partitions and let P satisfy $(x_0,x_1,\dots,x_k)=\mathrm{EP}(P)=\mathrm{EP}(P_f)\cup\mathrm{EP}(P_g)$. Then

$$A_+(f,P_f)+A_+(g,P_g)\geq A_+(f,P)+A_+(g,P)\geq A_+(f+g,P)\geq I_+(f+g).$$

We then have

$$I_+(f+g)\leq A_+(f,P_f)+A_+(g,P_g)\quad\Longrightarrow\quad I_+(f+g)\leq I_+(f)+I_+(g).$$

In like fashion we obtain the estimate

$$I_-(f+g)\geq I_-(f)+I_-(g).$$

Combining this gives

$$I_-(f)+I_-(g)\leq I_-(f+g)=I_+(f+g)\leq I_+(f)+I_+(g),$$

which implies equality of these four terms since $I_-(f)=I_+(f)$ and $I_-(g)=I_+(g)$. This gives this part of the result when I is compact. The result follows for general intervals from the definition of the Riemann integral for such intervals, and by applying Proposition 2.3.23.

(ii) As in part (i), the result will follow if we can prove it when I is compact. When $c=0$ the result is trivial, so suppose that $c\neq 0$. First consider the case $c>0$. For $\epsilon\in\mathbb{R}_{>0}$ let P be a partition for which $A_+(f,P)-A_-(f,P)<\frac{\epsilon}{c}$. Since $A_-(cf,P)=cA_-(f,P)$ and $A_+(cf,P)=cA_+(f,P)$ (as is easily checked), we have $A_+(cf,P)-A_-(cf,P)<\epsilon$, showing that cf is Riemann integrable. The equalities $A_-(cf,P)=cA_-(f,P)$ and $A_+(cf,P)=cA_+(f,P)$ then directly imply that $I_-(cf)=cI_-(f)$ and $I_+(cf)=cI_+(f)$, giving the result for $c>0$. For $c<0$ a similar argument holds, but asking that P be a partition for which $A_+(f,P)-A_-(f,P)<-\frac{\epsilon}{c}$.

(iii) First let us show that if I is compact then f^2 is Riemann integrable if f is Riemann integrable. This, however, follows from Proposition 3.4.23 by taking $g\colon I\to\mathbb{R}$ to be $g(x)=x^2$. To show that a general product fg of Riemann integrable functions on a compact interval is Riemann integrable, we note that

$$fg=\tfrac{1}{2}((f+g)^2-f^2-g^2).$$

By part (i) and using the fact that the square of a Riemann integrable function is Riemann integrable, the function on the right is Riemann integrable, so giving the result.

(iv) That $\frac{1}{g}$ is Riemann integrable follows from Proposition 3.4.23 by taking $g\colon I\to\mathbb{R}$ to be $g(x)=\frac{1}{x}$. ■

In parts (iii) and (iv) we asked that the interval be compact. It is simple to find counterexamples which indicate that compactness of the interval is generally necessary (see Exercise 3.4.3).

We now consider the relationship between composition and Riemann integration.

3.4.23 Proposition (Function composition and the Riemann integral) *If* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *is a compact interval, if* $\mathrm{f}\colon [\mathrm{a}, \mathrm{b}] \to \mathbb{R}$ *is a Riemann integrable function satisfying* $\mathrm{image}(\mathrm{f}) \subseteq [\mathrm{c}, \mathrm{d}]$*, and if* $\mathrm{g}\colon [\mathrm{c}, \mathrm{d}] \to \mathbb{R}$ *is continuous, then* $\mathrm{g} \circ \mathrm{f}$ *is Riemann integrable.*

Proof Denote $M = \sup\{|g(y)| \mid y \in [c, d]\}$. Let $\epsilon \in \mathbb{R}_{>0}$ and write $\epsilon' = \frac{\epsilon}{2M+d-c}$. Since g is uniformly continuous by the Heine–Cantor Theorem, let $\delta \in \mathbb{R}$ be chosen such that $0 < \delta < \epsilon'$ and such that, $|y_1 - y_2| < \delta$ implies that $|g(y_1) - g(y_2)| < \epsilon'$. Then choose a partition P of $[a, b]$ such that $A_+(f, P) - A_-(f, P) < \delta^2$. Let $(x_0, x_1, \dots, x_k)$ be the endpoints of P and define

$$\begin{aligned} A &= \{j \in \{1, \dots, k\} \mid \ \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\} < \delta\}, \\ B &= \{j \in \{1, \dots, k\} \mid \ \sup\{f(x) \mid x \in [x_{j-1}, x_j]\} - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\} \geq \delta\}. \end{aligned}$$

For $j \in A$ we have $|f(\xi_1) - f(\xi_2)| < \delta$ for every $\xi_1, \xi_2 \in [x_{j-1}, x_j]$ which implies that $|g \circ f(\xi_1) - g \circ f(\xi_2)| < \epsilon'$ for every $\xi_1, \xi_2 \in [x_{j-1}, x_j]$. For $j \in B$ we have

$$\begin{aligned} \delta \sum_{j \in B} (x_j - x_{j-1}) \leq & \sum_{j \in B} \Big(\sup\{f(x) \mid x \in [x_{j-1}, x_j]\} \\ & - \inf\{f(x) \mid x \in [x_{j-1}, x_j]\} \Big) (x_j - x_{j-1}) \\ \leq & \, A_+(f, P) - A_-(f, P) < \delta^2. \end{aligned}$$

Therefore we conclude that

$$\sum_{j \in B} (x_j - x_{j-1}) \leq \epsilon'.$$

Thus

$$\begin{aligned} A_+(g \circ f, P) - A_-(g \circ f, P) = & \sum_{j=1}^{k} \Big(\sup\{g \circ f(x) \mid x \in [x_{j-1}, x_j]\} \\ & - \inf\{g \circ f(x) \mid x \in [x_{j-1}, x_j]\} \Big) (x_j - x_{j-1}) \\ = & \sum_{j \in A} \Big(\sup\{g \circ f(x) \mid x \in [x_{j-1}, x_j]\} \\ & - \inf\{g \circ f(x) \mid x \in [x_{j-1}, x_j]\} \Big) (x_j - x_{j-1}) \\ & + \sum_{j \in B} \Big(\sup\{g \circ f(x) \mid x \in [x_{j-1}, x_j]\} \\ & - \inf\{g \circ f(x) \mid x \in [x_{j-1}, x_j]\} \Big) (x_j - x_{j-1}) \\ < & \, \epsilon'(d - c) + 2\epsilon' M < \epsilon, \end{aligned}$$

giving the result by Theorem 3.4.9. ■

The Riemann integral also has the expected properties relative to the partial order and the absolute value function on $\mathbb{R}$.

3.4.24 Proposition (Riemann integral and total order on $\mathbb{R}$) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval and let* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *be (conditionally) Riemann integrable functions for which* $\mathrm{f}(\mathrm{x}) \leq \mathrm{g}(\mathrm{x})$ *for each* $\mathrm{x} \in \mathrm{I}$*. Then*

$$(\mathrm{C}) \int_{\mathrm{I}} \mathrm{f}(\mathrm{x}) \, \mathrm{dx} \leq (\mathrm{C}) \int_{\mathrm{I}} \mathrm{g}(\mathrm{x}) \, \mathrm{dx}.$$

Proof Note that by part (i) of Proposition 3.4.22 it suffices to take $f = 0$ and then show that $\int_I g(x)\,dx \geq 0$. In the case where $I = [a,b]$ we have

$$\int_a^b g(x)\,dx \geq (b-a)\inf\{g(x) \mid x \in [a,b]\} \geq 0,$$

which gives the result in this case. The result for general intervals follows from the definition, and the fact the a limit of nonnegative numbers is nonnegative. ■

3.4.25 Proposition (Riemann integral and absolute value on $\mathbb{R}$) *Let* I *be an interval, let* $f\colon I \to \mathbb{R}$*, and define* $|f|\colon I \to \mathbb{R}$ *by* $|f|(x) = |f(x)|$*. Then the following statements hold:*

(i) if f *is Riemann integrable then* $|f|$ *is Riemann integrable;*

(ii) if I *is compact and if* f *is conditionally Riemann integrable then* $|f|$ *is conditionally Riemann integrable.*

Moreover, if the hypotheses of either part hold then

$$\left|\int_I f(x)\,dx\right| \leq \int_I |f|(x)\,dx.$$

Proof (i) If f is Riemann integrable then f_+ and f_- are Riemann integrable. Since $|f| = f_+ + f_-$ it follows from Proposition 3.4.22 that $|f|$ is Riemann integrable.

(ii) When I is compact, the statement follows since conditional Riemann integrability is equivalent to Riemann integrability.

The inequality in the statement of the proposition follows from Proposition 3.4.24 since $f(x) \leq |f(x)|$ for all $x \in I$. ■

We comment that the preceding result is, in fact, not true if one removes the condition that I be compact. We also comment that the converse of the result is false, in that the Riemann integrability of $|f|$ does not imply the Riemann integrability of f. The reader is asked to sort this out in Exercise 3.4.4.

The Riemann integral also behaves well upon breaking an interval into two intervals that are disjoint except for a common endpoint.

3.4.26 Proposition (Breaking the Riemann integral in two) *Let* $I \subseteq \mathbb{R}$ *be an interval and let* $I = I_1 \cup I_2$*, where* $I_1 \cap I_2 = \{c\}$*, where* c *is the right endpoint of* I_1 *and the left endpoint of* I_2*. Then* $f\colon I \to \mathbb{R}$ *is (conditionally) Riemann integrable if and only if* $f|I_1$ *and* $f|I_2$ *are (conditionally) Riemann integrable. Furthermore, we have*

$$(C)\int_I f(x)\,dx = (C)\int_{I_1} f(x)\,dx + (C)\int_{I_2} f(x)\,dx.$$

Proof We first consider the case where $I_1 = [a,c]$ and $I_2 = [c,b]$.

Let us suppose that f is Riemann integrable and let $(x_0, x_1, \dots, x_k)$ be endpoints of a partition of $[a,b]$ for which $A_+(f,P) - A_-(f,P) < \epsilon$. If $c \in (x_0, x_1, \dots, x_k)$, say $c = x_j$, then we have

$$A_-(f,P) = A_-(f|I_1,P_1) + A_-(f|I_2,P_2), \quad A_+(f,P) = A_+(f|I_1,P_1) + A_+(f|I_2,P_2),$$

where $\mathrm{EP}(P_1) = (x_0, x_1, \dots, x_j)$ are the endpoints of a partition of $[a, c]$ and $\mathrm{EP}(P_2) = (x_j, \dots, x_k)$ is a partition of $[c, b]$. From this we directly deduce that

$$A_+(f|I_1, P_1) - A_-(f|I_1, P_1) < \epsilon, \quad A_+(f|I_2, P_2) - A_-(f|I_2, P_2) < \epsilon. \tag{3.14}$$

If c is not an endpoint of P, then one can construct a new partition P' of $[a, b]$ with c as an extra endpoint. By Lemma 1 of Theorem 3.4.9 we have $A_+(f, P') - A_-(f, P') < \epsilon$. The argument then proceeds as above to show that (3.14) holds. Thus $f|I_1$ and $f|I_2$ are Riemann integrable by Theorem 3.4.9.

To prove the equality of the integrals in the statement of the proposition, we proceed as follows. Let P_1 and P_2 be partitions of I_1 and I_2, respectively. From these construct a partition $P(P_1, P_2)$ of I by asking that $\mathrm{EP}(P(P_1, P_2)) = \mathrm{EP}(P_1) \cup \mathrm{EP}(P_2)$. Then

$$A_+(f|I_1, P_1) + A_+(f|I_2, P_2) = A_+(f, P(P_1, P_2)).$$

Thus

$$\begin{aligned}\inf\{A_+(f|I_1, P_1) \mid P_1 \in \mathrm{Part}(I_1)\} + \inf\{A_+(f|I_2, P_2) \mid P_2 \in \mathrm{Part}(I_2)\}& \\ \geq \inf\{A_+(f, P) \mid P \in \mathrm{Part}(I)\}.&\end{aligned} \tag{3.15}$$

Now let P be a partition of I and construct partitions $P_1(P)$ and $P_2(P)$ of I_1 and I_2 respectively by adding defining, if necessary, a new partition P' of I with c as the (say) jth endpoint, and then defining $P_1(P)$ such that $\mathrm{EP}(P_1(P))$ are the first $j + 1$ endpoints of P' and then defining $P_2(P)$ such that $\mathrm{EP}(P_2(P))$ are the last $k - j$ endpoints of P'. By Lemma 1 of Theorem 3.4.9 we then have

$$A_+(f, P) \geq A_+(f, P') = A_+(f|I_1, P_1(P)) + A_+(f|I_2, P_2(P)).$$

This gives

$$\begin{aligned}&\inf\{A_+(f, P) \mid P \in \mathrm{Part}(I)\}\\ &\qquad \geq \inf\{A_+(f|I_1, P_1) \mid P_1 \in \mathrm{Part}(I_1)\} + \inf\{A_+(f|I_2, P_2) \mid P_2 \in \mathrm{Part}(I_2)\}.\end{aligned}$$

Combining this with (3.15) gives

$$\begin{aligned}&\inf\{A_+(f, P) \mid P \in \mathrm{Part}(I)\}\\ &\qquad = \inf\{A_+(f|I_1, P_1) \mid P_1 \in \mathrm{Part}(I_1)\} + \inf\{A_+(f|I_2, P_2) \mid P_2 \in \mathrm{Part}(I_2)\},\end{aligned}$$

which is exactly the desired result.

The result for a general interval follows from the general definition of the Riemann integral, and from Proposition 2.3.23. ■

The next result gives a useful tool for evaluating integrals, as well as a being a result of some fundamental importance.

3.4.27 Proposition (Change of variables for the Riemann integral) *Let* $[a, b]$ *be a compact interval and let* $u\colon [a, b] \to \mathbb{R}$ *be differentiable with* u' *Riemann integrable. Suppose that* $\mathrm{image}(u) \subseteq [c, d]$ *and that* $f\colon [c, d] \to \mathbb{R}$ *is Riemann integrable and that* $f = F'$ *for some differentiable function* $F\colon [c, d] \to \mathbb{R}$*. Then*

$$\int_a^b f \circ u(x) u'(x)\, dx = \int_{u(a)}^{u(b)} f(y)\, dy.$$

Proof Let $G\colon [a,b] \to \mathbb{R}$ be defined by $G = F \circ u$. Then $G' = (f \circ u)u'$ by the Chain Rule. Moreover, G' is Riemann integrable by Propositions 3.4.22 and 3.4.23. Thus, twice using Theorem 3.4.30 below,

$$\int_a^b f \circ u(x)u'(x)\,\mathrm{d}x = G(b) - G(a) = F \circ u(b) - F \circ u(a) = \int_{u(a)}^{u(b)} f(y)\,\mathrm{d}y,$$

as desired. ■

As a final result in this section, we prove the extremely valuable integration by parts formula.

3.4.28 Proposition (Integration by parts for the Riemann integral) *If* $[\mathrm{a},\mathrm{b}]$ *is a compact interval and if* $\mathrm{f},\mathrm{g}\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}$ *are differentiable functions with* f' *and* g' *Riemann integrable, then*

$$\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}(\mathrm{x})\mathrm{g}'(\mathrm{x})\,\mathrm{d}\mathrm{x} + \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}'(\mathrm{x})\mathrm{g}(\mathrm{x})\,\mathrm{d}\mathrm{x} = \mathrm{f}(\mathrm{b})\mathrm{g}(\mathrm{b}) - \mathrm{f}(\mathrm{a})\mathrm{g}(\mathrm{a}).$$

Proof By Proposition 3.2.10 it holds that fg is differentiable and that $(fg)' = f'g + fg'$. Thus, by Proposition 3.4.22, fg is differentiable with Riemann integrable derivative. Therefore, by Theorem 3.4.30 below,

$$\int_a^b (fg)(x)\,\mathrm{d}x = f(b)g(b) - f(a)g(a),$$

and the result follows directly from the formula for the product rule. ■

3.4.6 The Fundamental Theorem of Calculus and the Mean Value Theorems

In this section we begin to explore the sense in which differentiation and integration are inverses of one another. This is, in actuality, and somewhat in contrast to the manner in which one considers this question in introductory calculus courses, a quite complicated matter. Indeed, we will not fully answer this question until Section 5.9.7, after we have some knowledge of the Lebesgue integral. Nevertheless, in this section we give some simple results, and some examples which illustrate the value and the limitations of these results. We also present the Mean Value Theorems for integrals.

The following language is often used in conjunction with the Fundamental Theorem of Calculus.

3.4.29 Definition (Primitive) If $I \subseteq \mathbb{R}$ is an interval and if $f\colon I \to \mathbb{R}$ is a function, a ***primitive*** for f is a function $F\colon I \to \mathbb{R}$ such that $F' = f$. •

Note that primitives are not unique since if one adds a constant to a primitive, the resulting function is again a primitive.

The basic result of this section is the following.

3.4.30 Theorem (Fundamental Theorem of Calculus for Riemann integrals) *For a compact interval* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$*, the following statements hold:*

(i) *if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is Riemann integrable with primitive* $\mathrm{F}\colon \mathrm{I} \to \mathbb{R}$*, then*

$$\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}(\mathrm{x})\,\mathrm{d}\mathrm{x} = \mathrm{F}(\mathrm{b}) - \mathrm{F}(\mathrm{a});$$

(ii) *if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is Riemann integrable, and if* $\mathrm{F}\colon \mathrm{I} \to \mathbb{R}$ *is defined by*

$$\mathrm{F}(\mathrm{x}) = \int_{\mathrm{a}}^{\mathrm{x}} \mathrm{f}(\xi)\,\mathrm{d}\xi,$$

then

(a) F *is continuous and*

(b) *at each point* $\mathrm{x} \in \mathrm{I}$ *for which* f *is continuous,* F *is differentiable and* $\mathrm{F}'(\mathrm{x}) = \mathrm{f}(\mathrm{x})$*.*

Proof (i) Let $(P_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of partitions for which $\lim_{j\to\infty}|P_j| = 0$. Denote by $(x_{j,0}, x_{j,1}, \dots, x_{j,k_j})$ the endpoints of P_j, $j \in \mathbb{Z}_{>0}$. By the Mean Value Theorem, for each $j \in \mathbb{Z}_{>0}$ and for each $r \in \{1, \dots, k_r\}$, there exists $\xi_{j,r} \in [x_{j,r-1}, x_{j,r}]$ such that $F(x_{j,r}) - F(x_{j,r-1}) = f(\xi_{j,r})(x_{j,r} - x_{j,r-1})$. Since f is Riemann integrable we have

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x &= \lim_{j\to\infty}\sum_{r=1}^{k_j} f(\xi_{j,r})(x_{j,r} - x_{j,r-1})\\ &= \lim_{j\to\infty}\sum_{r=1}^{k_j}(F(x_{j,r}) - F(x_{j,r-1}))\\ &= \lim_{j\to\infty}(F(b) - F(a)) = F(b) - F(a),\end{aligned}$$

as desired.

(ii) Let $x \in (a, b)$ and note that, for h sufficiently small,

$$F(x+h) - F(x) = \int_x^{x+h} f(\xi)\,\mathrm{d}\xi,$$

using Proposition 3.4.26. By Proposition 3.4.24 it follows that

$$h\inf\{f(y) \mid y \in [a,b]\} \le \int_x^{x+h} f(\xi)\,\mathrm{d}\xi \le h\sup\{f(y) \mid y \in [a,b]\},$$

provided that $h > 0$. This shows that

$$\lim_{h\downarrow 0}\int_x^{x+h} f(\xi)\,\mathrm{d}\xi = 0.$$

A similar argument can be fashioned for the case when $h < 0$ to show also that

$$\lim_{h\uparrow 0}\int_x^{x+h} f(\xi)\,\mathrm{d}\xi = 0,$$

so showing that F is continuous at point in (a, b). A slight modification to this argument shows that F is also continuous at a and b.

Now suppose that f is continuous at x. Let $h > 0$. Again using Proposition 3.4.24 we have

$$\begin{aligned} h\inf\{f(y) \mid y \in [x, x+h]\} &\le \int_x^{x+h} f(\xi)\,\mathrm{d}\xi \le h\sup\{f(y) \mid y \in [x, x+h]\} \\ \implies \quad \inf\{f(y) \mid y \in [x, x+h]\} &\le \frac{F(x+h) - F(x)}{h} \le \sup\{f(y) \mid y \in [x, x+h]\}. \end{aligned}$$

Continuity of f at x gives

$$\lim_{h\downarrow 0}\inf\{f(y) \mid y \in [x, x+h]\} = f(x), \quad \lim_{h\downarrow 0}\sup\{f(y) \mid y \in [x, x+h]\} = f(x).$$

Therefore,

$$\lim_{h\downarrow 0}\frac{F(x+h) - F(x)}{h} = f(x).$$

A similar argument can be made for $h < 0$ to give

$$\lim_{h\uparrow 0}\frac{F(x+h) - F(x)}{h} = f(x),$$

so proving this part of the theorem. ■

Let us give some examples that illustrate what the Fundamental Theorem of Calculus says and does not say.

3.4.31 Examples (Fundamental Theorem of Calculus)

1. Let $I = [0, 1]$ and define $f: I \to \mathbb{R}$ by

$$f(x) = \begin{cases} x, & x \in [0, \frac{1}{2}], \\ 1 - x, & x \in (\frac{1}{2}, 1]. \end{cases}$$

Then

$$F(x) \triangleq \int_0^x f(\xi)\,\mathrm{d}\xi = \begin{cases} \frac{1}{2}x^2, & x \in [0, \frac{1}{2}], \\ -\frac{1}{2}x^2 + x - \frac{1}{8}, & x \in (\frac{1}{2}, 1]. \end{cases}$$

Then, for any $x \in [a, b]$, we see that

$$\int_0^x f(\xi)\,\mathrm{d}\xi = F(x) - F(0).$$

This is consistent with part (i) of Theorem 3.4.30, whose hypotheses apply since f is continuous, and so Riemann integrable.

2. Let $I = [0, 1]$ and define $f: I \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}], \\ -1, & x \in (\frac{1}{2}, 1]. \end{cases}$$

Then

$$F(x) \triangleq \int_0^x f(\xi)\,\mathrm{d}\xi = \begin{cases} x, & x \in [0, \frac{1}{2}], \\ 1 - x, & x \in (\frac{1}{2}, 1]. \end{cases}$$

Then, for any $x \in [a, b]$, we see that

$$\int_0^x f(\xi)\,\mathrm{d}\xi = F(x) - F(0).$$

In this case, we have the conclusions of part (i) of Theorem 3.4.30, and indeed the hypotheses hold, since f is Riemann integrable.

3. Let I and f be as in Example 1 above. Then f is Riemann integrable, and we see that F is continuous, as per part (ii) of Theorem 3.4.30, and that F is differentiable, also as per part (ii) of Theorem 3.4.30.
4. Let I and f be as in Example 2 above. Then f is Riemann integrable, and we see that F is continuous, as per part (ii) of Theorem 3.4.30. However, f is not continuous at $x = \frac{1}{2}$, and we see that, correspondingly, F is not differentiable at $x = \frac{1}{2}$.
5. The next example we consider is one with which, at this point, we can only be sketchy about the details. Consider the Cantor function $f_C\colon [0, 1] \to \mathbb{R}$ of Example 3.2.27. Note that f_C' is defined and equal to zero, except at points in the Cantor set C; thus except at points forming a set of measure zero. It will be clear when we discuss the Lebesgue integral in Section 5.9 that this ensures that $\int_0^x f_C'(\xi)\,\mathrm{d}\xi = 0$ for every $x \in [0, 1]$, where the integral in this case is the Lebesgue integral. (By defining f_C' arbitrarily on C, we can also use the Riemann integral by virtue of Theorem 3.4.11.) This shows that the conclusions of part (i) of Theorem 3.4.30 can fail to hold, even when the derivative of F is defined almost everywhere.
6. The last example we give is the most significant, in some sense, and is also the most complicated. The example we give is of a function $F\colon [0, 1] \to \mathbb{R}$ that is differentiable with bounded derivative, but whose derivative $f = F'$ is not Riemann integrable. Thus f possesses a primitive, but is not Riemann integrable.

   To define F, let $G\colon \mathbb{R}_{>0} \to \mathbb{R}$ be the function

$$G(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

   For $c > 0$ let $x_c > 0$ be defined by

$$x_c = \sup\{x \in \mathbb{R}_{>0} \mid G'(x) = 0,\ x \leq c\},$$

   and define $G_c\colon (0, c] \to \mathbb{R}$ by

$$G_c(x) = \begin{cases} G(x), & x \in (0, x_c], \\ G(x_c), & x \in (x_c, x]. \end{cases}$$

Now, for $\epsilon \in (0, \frac{1}{2})$, let $C_\epsilon \subseteq [0,1]$ be a fat Cantor set as constructed in Example 2.5.42. Define F as follows. If $x \in C_\epsilon$ we take $F(x) = 0$. If $x \notin C_\epsilon$, then, since C_ϵ is closed, by Proposition 2.5.6 x lies in some open interval, say (a, b). Then take $c = \frac{1}{2}(b-a)$ and define

$$F(x) = \begin{cases} G_c(x-a), & x \in (a, \frac{1}{2}(a+b)), \\ G_c(b-x), & x \in [\frac{1}{2}(a+b), b). \end{cases}$$

Note that $F|(a,b)$ is designed so that its derivative will oscillate wildly in the limit as the endpoints of (a,b) are approached, but be nicely behaved at all points in (a,b). This is, as we shall see, the key feature of F.

Let us record some properties of F in a sequence of lemmata.

1 Lemma *If* $\mathrm{x} \in \mathrm{C}_\epsilon$*, then* F *is differentiable at* x *and* $\mathrm{F}'(\mathrm{x}) = 0$.

Proof Let $y \in [0,1] \setminus \{x\}$. If $y \in C_\epsilon$ then

$$\frac{f(y) - f(x)}{y - x} = 0.$$

If $y \notin C_\epsilon$, then y must lie in an open interval, say (a,b). Let d be the endpoint of (a,b) nearest y and let $c = \frac{1}{2}(b-a)$. Then

$$\begin{aligned} \left|\frac{f(y) - f(x)}{y - x}\right| = \frac{f(y)}{y-x} \leq \frac{f(y)}{y-d} = \frac{G_c(|y-d|)}{y-d} \\ \leq \frac{|y-d|^2}{y-d} = |y-d| \leq |y-x|. \end{aligned}$$

Thus

$$\lim_{y \to x} \frac{f(y) - f(x)}{y - x} = 0,$$

giving the lemma. ▼

2 Lemma *If* $\mathrm{x} \notin \mathrm{C}_\epsilon$*, then* F *is differentiable at* x *and* $|\mathrm{F}'(\mathrm{x})| \leq 3$.

Proof By definition of F for points not in C_ϵ we have

$$|F'(x)| \leq \left|2y \sin \tfrac{1}{y} - \cos \tfrac{1}{y}\right| \leq 3,$$

for some $y \in [0,1]$. ▼

3 Lemma $\mathrm{C}_\epsilon \subseteq \mathrm{D}_{\mathrm{F}'}$.

Proof By construction of C_ϵ, if $x \in C_\epsilon$ then there exists a sequence $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ of open intervals in $[0,1] \setminus C_\epsilon$ having the property that $\lim_{j\to\infty} a_j = \lim_{j\to\infty} b_j = x$. Note that $\limsup_{y\downarrow 0} g'(y) = 1$. Therefore, by the definition of F on the open intervals (a_j, b_j), $j \in \mathbb{Z}_{>0}$, it holds that $\limsup_{y\downarrow a_j} F'(y) = \limsup_{y\uparrow b_j} F'(y) = 1$. Therefore, $\limsup_{y\to x} F'(y) = 1$. Since $F'(x) = 0$, it follows that F' is discontinuous at x. ▼

Since F' is discontinuous at all points in C_ϵ, and since C_ϵ does not have measure zero, it follows from Theorem 3.4.11 that F' is not Riemann integrable. Therefore, the function $f = F'$ possesses a primitive, namely F, but is not Riemann integrable. •

Finally we state two results that, like the Mean Value Theorem for differentiable functions, relate the integral to the values of a function.

3.4.32 Proposition (First Mean Value Theorem for Riemann integrals) *Let* $[\mathrm{a},\mathrm{b}]$ *be a compact interval and let* $\mathrm{f},\mathrm{g}\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}$ *be functions with* f *continuous and with* g *nonnegative and Riemann integrable. Then there exists* $\mathrm{c} \in [\mathrm{a},\mathrm{b}]$ *such that*

$$\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}(\mathrm{x})\mathrm{g}(\mathrm{x})\,\mathrm{dx} = \mathrm{f}(\mathrm{c}) \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{g}(\mathrm{x})\,\mathrm{dx}$$

Proof Let

$$m = \inf\{f(x) \mid x \in [a,b]\}, \quad M = \sup\{f(x) \mid x \in [a,b]\}.$$

Since g is nonnegative we have

$$mg(x) \leq f(x)g(x) \leq Mg(x), \qquad x \in [a,b],$$

from which we deduce that

$$m\int_a^b g(x)\,\mathrm{d}x \leq \int_a^b f(x)g(x)\,\mathrm{d}x \leq M\int_a^b g(x)\,\mathrm{d}x.$$

Continuity of f and the Intermediate Value Theorem gives $c \in [a,b]$ such that the result holds. ■

3.4.33 Proposition (Second Mean Value Theorem for Riemann integrals) *Let* $[\mathrm{a},\mathrm{b}]$ *be a compact interval and let* $\mathrm{f},\mathrm{g}\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}$ *be functions with*

(i) g *Riemann integrable and having the property that there exists* G *such that* $\mathrm{g} = \mathrm{G}'$, *and*

(ii) f *differentiable with Riemann integrable, nonnegative derivative.*

Then there exists $\mathrm{c} \in [\mathrm{a},\mathrm{b}]$ *so that*

$$\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}(\mathrm{x})\mathrm{g}(\mathrm{x})\,\mathrm{dx} = \mathrm{f}(\mathrm{a}) \int_{\mathrm{a}}^{\mathrm{c}} \mathrm{g}(\mathrm{x})\,\mathrm{dx} + \mathrm{f}(\mathrm{b}) \int_{\mathrm{c}}^{\mathrm{b}} \mathrm{g}(\mathrm{x})\,\mathrm{dx}.$$

Proof Without loss of generality we may suppose that

$$G(x) = \int_a^x g(\xi)\,\mathrm{d}\xi,$$

since all we require is that $G' = g$. We then compute

$$\begin{aligned}\int_a^b f(x)g(x)\,\mathrm{d}x = \int_a^b f(x)G'(x)\,\mathrm{d}x &= f(b)G(b) - \int_a^b f'(x)G(x)\,\mathrm{d}x \\ &= f(b)G(b) - G(c)\int_a^b f'(x)\,\mathrm{d}x,\end{aligned}$$

for some $c \in [a, b]$, using integration by parts and Proposition 3.4.32. Now using Theorem 3.4.30,

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(b)G(b) - G(c)(f(b) - f(a)),$$

which gives the desired result after using the definition of G and after some rearrangement. ■

3.4.7 The Cauchy principal value

In Example 3.4.17 we explored some of the nuances of the improper Riemann integral. There we saw that for integrals that are defined using limits, one often needs to make the definitions in a particular way. The principal value integral is intended to relax this, and enable one to have a meaningful notion of the integral in cases where otherwise one might not. To motivate our discussion we consider an example.

3.4.34 Example Let $I = [-1, 2]$ and consider the function $f: I \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{x}, & x \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

This function has a singularity at $x = 0$, and the integral $\int_{-1}^{2} f(x)\,\mathrm{d}x$ is actually divergent. However, for $\epsilon \in \mathbb{R}_{>0}$ note that

$$\int_{-1}^{-\epsilon} \frac{1}{x}\,\mathrm{d}x + \int_{\epsilon}^{2} \frac{1}{x}\,\mathrm{d}x = -\log x|_{\epsilon}^{1} + \log x|_{\epsilon}^{2} = \log 2.$$

Thus we can devise a way around the singularity in this case, the reason being that the singular behaviour of the function on either side of the function "cancels" that on the other side. •

With this as motivation, we give a definition.

3.4.35 Definition (Cauchy principal value) Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}$ be a function. Denote $a = \inf I$ and $b = \sup I$, allowing that $a = -\infty$ and $b = \infty$.

(i) If, for $x_0 \in \operatorname{int}(I)$, there exists $\epsilon_0 \in \mathbb{R}_{>0}$ such that the functions $f|(a, x_0 - \epsilon]$ and $f|[x_0 + \epsilon, b)$ are Riemann integrable for all $\epsilon \in (0, \epsilon_0]$, then the ***Cauchy principal value*** for f is defined by

$$\mathrm{pv}\int_I f(x)\,\mathrm{d}x = \lim_{\epsilon\to 0}\Big(\int_a^{x_0-\epsilon} f(x)\,\mathrm{d}x + \int_{x_0+\epsilon}^{b} f(x)\,\mathrm{d}x\Big).$$

(ii) If $a = -\infty$ and $b = \infty$ and if for each $R \in \mathbb{R}_{>0}$ the function $f|[-R, R]$ is Riemann integrable, then the ***Cauchy principal value*** for f is defined by

$$\mathrm{pv}\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \lim_{R\to\infty}\int_{-R}^{R} f(x)\,\mathrm{d}x.$$ •

3.4.36 Remarks

1. If f is Riemann integrable on I then the Cauchy principal value is equal to the Riemann integral.
2. The Cauchy principal value is allowed to be infinite by the preceding definition, as the following examples will show.
3. It is not standard to define the Cauchy principal value in part (ii) of the definition. In many texts where the Cauchy principal value is spoken of, it is part (i) that is being used. However, we will find the definition from part (ii) useful. •

3.4.37 Examples (Cauchy principal value)

1. For the example of Example 3.4.34 we have

$$\mathrm{pv}\int_{-1}^{2} \frac{1}{x}\,\mathrm{d}x = \log 2.$$

2. For $I = \mathbb{R}$ and $f(x) = x(1 + x^2)^{-1}$ we have

$$\mathrm{pv}\int_{-\infty}^{\infty} \frac{x}{1+x^2}\,\mathrm{d}x = \lim_{R\to\infty}\int_{-R}^{R} \frac{x}{1+x^2}\,\mathrm{d}x = \lim_{R\to\infty}\Big(\frac{1}{2}\log(1+R^2) - \frac{1}{2}\log(1+R^2)\Big) = 0.$$

Note that in Example 3.4.17–4 we showed that this function was not Riemann integrable.

3. Next we consider $I = \mathbb{R}$ and $f(x) = |x|(1 + x^2)$. In this case we compute

$$\mathrm{pv}\int_{-\infty}^{\infty} \frac{|x|}{1+x^2}\,\mathrm{d}x = \lim_{R\to\infty}\int_{-R}^{R} \frac{|x|}{1+x^2}\,\mathrm{d}x = \lim_{R\to\infty}\Big(\frac{1}{2}\log(1+R^2) + \frac{1}{2}\log(1+R^2)\Big) = \infty.$$

We see then that there is no reason why the Cauchy principal value may not be infinite. •

3.4.8 Notes

The definition we give for the Riemann integral is actually that used by Darboux, and the condition given in part (iii) of Theorem 3.4.9 is the original definition of Riemann. What Darboux showed was that the two definitions are equivalent. It is not uncommon to instead use the Darboux definition as the standard definition because, unlike the definition of Riemann, it does not rely on an arbitrary selection of a point from each of the intervals forming a partition.

Exercises

3.4.1 Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \mathbb{R}$ be a function that is Riemann integrable and satisfies $f(x) \geq 0$ for all $x \in I$. Show that $\int_I f(x)\,\mathrm{d}x \geq 0$.

3.4.2 Let $I \subseteq \mathbb{R}$ be an interval, let $f, g\colon I \to \mathbb{R}$ be functions, and define $D_{f,g} = \{x \in I \mid f(x) \neq g(x)\}$.

(a) Show that, if $D_{f,g}$ is finite and f is Riemann integrable, then g is Riemann integrable and $\int_I f(x)\,\mathrm{d}x = \int_I g(x)\,\mathrm{d}x$.

(b) Is it true that, if $D_{f,g}$ is countable and f is Riemann integrable, then g is Riemann integrable and $\int_I f(x)\,\mathrm{d}x = \int_I g(x)\,\mathrm{d}x$? If it is true, give a proof; if it is not true, give a counterexample.

3.4.3 Do the following:

(a) find an interval I and functions $f, g\colon I \to \mathbb{R}$ such that f and g are both Riemann integrable, but fg is not Riemann integrable;

(b) find an interval I and functions $f, g\colon I \to \mathbb{R}$ such that f and g are both Riemann integrable, but $g \circ f$ is not Riemann integrable.

3.4.4 Do the following:

(a) find an interval I and a conditionally Riemann integrable function $f\colon I \to \mathbb{R}$ such that $|f|$ is not Riemann integrable;

(b) find a function $f\colon [0, 1] \to \mathbb{R}$ such that $|f|$ is Riemann integrable, but f is not Riemann integrable.

3.4.5 Show that, if $f\colon [a, b] \to \mathbb{R}$ is continuous, then there exists $c \in [a, b]$ such that

$$\int_a^b f(x)\,\mathrm{d}x = f(c)(b - a).$$

Section 3.5

Sequences and series of ℝ-valued functions

In this section we present for the first time the important topic of sequences and series of functions and their convergence. One of the reasons why convergence of sequences of functions is important is that is allows us to classify sets of functions. The idea of classifying sets of functions according to their possessing certain properties leads to the general idea of a "function space." Function spaces are important to understand when developing any systematic theory dealing with functions, since sets of general functions are simply too unstructured to allow much useful to be said. On the other hand, if one restricts the set of functions in the wrong way (e.g., by asking that they all be continuous), then one can end of with a framework with unpleasant properties. But this is getting a little ahead of the issue directly at hand, which is to consider convergence of sequences of functions.

Do I need to read this section? The material in this section is basic, particularly the concepts of pointwise convergence and uniform convergence and the distinction between them. However, it is possible to avoid reading this section until the material becomes necessary, as it will in Chapters 10, 11, 12, and 13, for example. •

3.5.1 Pointwise convergent sequences

The first type of convergence we deal with is probably what a typical first-year student, at least the rare one who understood convergence for summations of numbers, would proffer as a good candidate for convergence. As we shall see, it often leaves something to be desired.

In the discussion of pointwise convergence, one needs no assumptions on the character of the functions, as one is essentially talking about convergence of numbers.

3.5.1 Definition (Pointwise convergence of sequences) Let $I \subseteq \mathbb{R}$ be an interval and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of ℝ-valued functions on I.

(i) The sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ ***converges pointwise*** to a function $f\colon I \to \mathbb{R}$ if, for each $x \in I$ and for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \epsilon$ provided that $j \geq N$.

(ii) The function f in the preceding part of the definition is the ***limit function*** for the sequence.

(iii) The sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ is ***pointwise Cauchy*** if, for each $x \in I$ and for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f_k(x)| < \epsilon$ provided that $j, k \geq N$. •

Let us immediately establish the equivalence of pointwise convergent and pointwise Cauchy sequences. As is clear in the proof of the following result, the key fact is completeness of ℝ.

3.5.2 Theorem (Pointwise convergent equals pointwise Cauchy) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a sequence of* $\mathbb{R}$*-valued functions on* I *then the following statements are equivalent:*

(i) *there exists a function* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *such that* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges pointwise to* f*;*

(ii) $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is pointwise Cauchy.*

Proof This merely follows from the following facts.

1. If the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ converges to $f(x)$ then the sequence is Cauchy by Proposition 2.3.3.
2. If the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ is Cauchy then there exists a number $f(x) \in \mathbb{R}$ such that $\lim_{j\to\infty} f_j(x) = f(x)$ by Theorem 2.3.5. ■

Based on the preceding theorem we shall switch freely between the notions of pointwise convergent and pointwise Cauchy sequences of functions.

Pointwise convergence is essentially the most natural form of convergence for a sequence of functions in that it depends in a trivial way on the basic notion of convergence of sequences in $\mathbb{R}$. However, as we shall see later in this section, and in Chapters 6 and **??**, other forms of convergence of often more useful.

3.5.3 Example (Pointwise convergence) Consider the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ of $\mathbb{R}$-valued functions defined on $[0,1]$ by

$$f_j(x) = \begin{cases} 1, & x \in [0, \frac{1}{j}], \\ 0, & x \in (\frac{1}{j}, 1]. \end{cases}$$

Note that $f_j(0) = 1$ for every $j \in \mathbb{Z}_{>0}$, so that the sequence $(f_j(0))_{j\in\mathbb{Z}_{>0}}$ converges, trivially, to 1. For any $x_0 \in (0,1]$, provided that $j > x_0^{-1}$, then $f_j(x_0) = 0$. Thus $(f_j(x_0))_{j\in\mathbb{Z}_{>0}}$ converges, as a sequence of real numbers, to 0 for each $x_0 \in (0,1]$. Thus this sequence converges pointwise, and the limit function is

$$f(x) = \begin{cases} 1, & x = 0, \\ 0, & x \in (0,1]. \end{cases}$$

If N is the smallest natural number with the property that $N > x_0^{-1}$, then we observe, trivially, that this number does indeed depend on x_0. As x_0 gets closer and closer to 0 we have to wait longer and longer in the sequence $(f_j(x_0))_{j\in\mathbb{Z}_{>0}}$ for the arrival of zero. •

3.5.2 Uniformly convergent sequences

Let us first say what we mean by uniform convergence.

3.5.4 Definition (Uniform convergence of sequences) Let $I \subseteq \mathbb{R}$ be an interval and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of $\mathbb{R}$-valued functions on I.

(i) The sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ ***converges uniformly*** to a function $f\colon I \to \mathbb{R}$ if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \epsilon$ for all $x \in I$, provided that $j \geq N$.

(ii) The sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ is ***uniformly Cauchy*** if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f_k(x)| < \epsilon$ for all $x \in I$, provided that $j, k \geq N$. •

Let us immediately give the equivalence of the preceding notions of convergence.

3.5.5 Theorem (Uniformly convergent equals uniformly Cauchy) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$ *and a sequence of* $\mathbb{R}$*-valued functions* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *on* I *the following statements are equivalent:*

(i) there exists a function $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *such that* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges uniformly to* f*;*

(ii) $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is uniformly Cauchy.*

Proof First suppose that $(f_j)_{j\in\mathbb{Z}_{>0}}$ is uniformly Cauchy. Then, for each $x \in I$ the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ is Cauchy and so by Theorem 2.3.5 converges to a number that we denote by $f(x)$. This defines the function $f\colon I \to \mathbb{R}$ to which the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges pointwise. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N_1 \in \mathbb{Z}_{>0}$ have the property that $|f_j(x) - f_k(x)| < \frac{\epsilon}{2}$ for $j, k \geq N_1$ and for each $x \in I$. Now let $x \in I$ and let $N_2 \in \mathbb{Z}_{>0}$ have the property that $|f_k(x) - f(x)| < \frac{\epsilon}{2}$ for $k \geq N_2$. Then, for $j \geq N_1$, we compute

$$|f_j(x) - f(x)| \leq |f_j(x) - f_k(x)| + |f_k(x) - f(x)| < \epsilon,$$

where $k \geq \max\{N_1, N_2\}$, giving the first implication.

Now suppose that, for $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f(x)| < \epsilon$ for all $j \geq N$ and for all $x \in I$. Then, for $\epsilon \in \mathbb{R}_{>0}$ let $N \in \mathbb{Z}_{>0}$ satisfy $|f_j(x) - f(x)| < \frac{\epsilon}{2}$ for $j \geq N$ and $x \in I$. Then, for $j, k \geq N$ and for $x \in I$, we have

$$|f_j(x) - f_k(x)| \leq |f_j(x) - f(x)| + |f_k(x) - f(x)| < \epsilon,$$

giving the sequence as uniformly Cauchy. ■

Compare this definition to that for pointwise convergence. They sound similar, but there is a fundamental difference. For pointwise convergence, the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ is examined separately for convergence at each value of x. As a consequence of this, the value of N might depend on both ϵ and x. For uniform convergence, however, we ask that for a given ϵ, the convergence is tested over all of I. In Figure 3.11 we depict the idea behind uniform convergence. The distinction between uniform and pointwise convergence is subtle on a first encounter, and it is sometimes difficult to believe that pointwise convergence is possible without uniform convergence. However, this is indeed the case, and an example illustrates this readily.

3.5.6 Example (Uniform convergence) On $[0, 1]$ we consider the sequence of $\mathbb{R}$-valued functions defined by

$$f_j(x) = \begin{cases} 2jx, & x \in [0, \frac{1}{2j}], \\ -2jx + 2, & x \in (\frac{1}{2j}, \frac{1}{j}], \\ 0, & x \in (\frac{1}{j}, 1]. \end{cases}$$

In Figure 3.12 we graph f_j for $j \in \{1, 3, 10, 50\}$. The astute reader will see the point, but let's go through it just to make sure we see how this works.

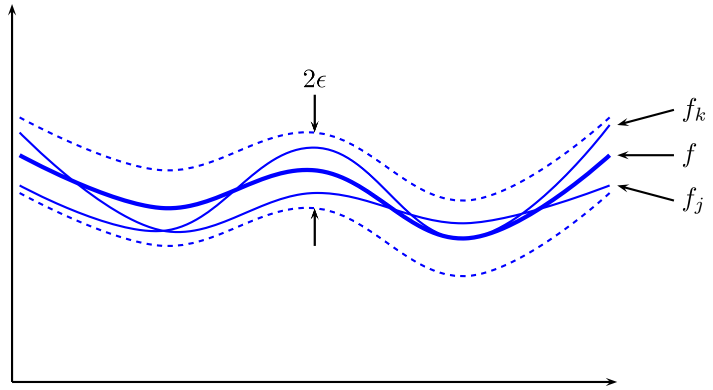


Figure 3.11 The idea behind uniform convergence

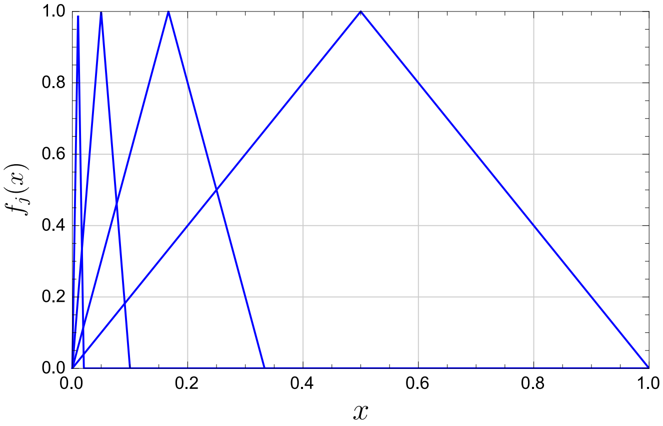


Figure 3.12 A sequence of functions converging pointwise, but not uniformly

First of all, we claim that the sequence converges pointwise to the limit function $f(x) = 0$, $x \in [0, 1]$. Since $f_j(0) = 0$ for all $j \in \mathbb{Z}_{>0}$, obviously the sequence converges to 0 at $x = 0$. For $x \in (0, 1]$, if $N \in \mathbb{Z}_{>0}$ satisfies $\frac{1}{N} < x$ then we have $f_j(x) = 0$ for $j \geq N$. Thus we do indeed have pointwise convergence.

We also claim that the sequence does not converge uniformly. Indeed, for any positive $\epsilon < 1$, we see that $f_j(\frac{1}{2j}) = 1 > \epsilon$ for every $j \in \mathbb{Z}_{>0}$. This prohibits our asserting the existence of $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f_k(x)| < \epsilon$ for every $x \in [0, 1]$, provided that $j, k \geq N$. Thus convergence is indeed not uniform. •

As we say, this is perhaps subtle, at least until one comes to grips with, after which point it makes perfect sense. You should not stop thinking about this until it makes perfect sense. If you overlook this distinction between pointwise and uniform convergence, you will be missing one of the most important topics in the theory of frequency representations of signals.

3.5.7 Remark (On "uniformly" again) In Remark 3.1.6 we made some comments on the notion of what is meant by "uniformly." Let us reinforce this here. In Definition 3.1.5 we introduced the notion of uniform continuity, which meant that the "δ" could be chosen so as to be valid on the entire domain. Here, with uniform convergence, the idea is that "N" can be chosen to be valid on the entire domain. Similar uses will occasionally be made of the word "uniformly" throughout the text, and it is hoped that the meaning should be clear from the context. •

Now we prove an important result concerning uniform convergence. The significance of this result is perhaps best recognised in a more general setting, such as that of Theorem **??**, where the idea of completeness is clear. However, even in the simple setting of our present discussion, the result is important enough.

3.5.8 Theorem (The uniform limit of bounded, continuous functions is bounded and continuous) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval with* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *a sequence of continuous bounded functions on* I *that converge uniformly. Then the limit function is continuous and bounded. In particular, a uniformly convergent sequence of continuous functions defined on a compact interval converges to a continuous limit function.*

Proof Let $x \in I$ define $f(x) = \lim_{j\to\infty} f_j(x)$. This pointwise limit exists since $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{R}$ (why?). We first claim that f is bounded. To see this, for $\epsilon \in \mathbb{R}_{>0}$, let $N \in \mathbb{Z}_{>0}$ have the property that $|f(x) - f_N(x)| < \epsilon$ for every $x \in I$. Then

$$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq \epsilon + \sup\{f_N(x) \mid x \in I\}.$$

Since the expression on the right is independent of x, this gives the desired boundedness of f.

Now we prove that the limit function f is continuous. Since $(f_j)_{j\in\mathbb{Z}_{>0}}$ is uniformly convergent, for any $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f(x)| < \frac{\epsilon}{3}$ for all $x \in I$ and $j \geq N$. Now fix $x_0 \in I$, and consider the $N \in \mathbb{Z}_{>0}$ just defined. By continuity of f_N, there exists $\delta \in \mathbb{R}_{>0}$ such that, if $x \in I$ satisfies $|x - x_0| < \delta$, then $|f_N(x) - f_N(x_0)| < \frac{\epsilon}{3}$. Then, for $x \in I$ satisfying $|x - x_0| < \delta$, we have

$$\begin{aligned}|f(x) - f(x_0)| &= |(f(x) - f_N(x)) + (f_N(x) - f_N(x_0)) + (f_N(x_0) - f(x_0))|\\ &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)|\\ &< \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon,\end{aligned}$$

where we have again used the triangle inequality. Since this argument is valid for any $x_0 \in I$, it follows that f is continuous. ■

Note that the hypothesis that the functions be bounded is essential for the conclusions to hold. As we shall see, the contrapositive of this result is often helpful. That is, it is useful to remember that if a sequence of continuous functions defined on a closed bounded interval converges to a *dis*continuous limit function, then the convergence is *not* uniform.

3.5.3 Dominated and bounded convergent sequences

Bounded convergence is a notion that is particularly useful when discussing convergence of function sequences on noncompact intervals.

3.5.9 Definition (Dominated and bounded convergence of sequences) Let $I \subseteq \mathbb{R}$ be an interval and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of $\mathbb{R}$-valued functions on I. For a function $g\colon I \to \mathbb{R}_{>0}$, the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ ***converges dominated by* g** if

(i) $f_j(x) \leq g(x)$ for every $j \in \mathbb{Z}_{>0}$ and for every $x \in I$ and

(ii) if, for each $x \in I$ and for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f_k(x)| < \epsilon$ for $j, k \geq N$.

If, moreover, g is a constant function, then a sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ that converges dominated by g ***converges boundedly***. •

It is clear that dominated convergence implies pointwise convergence. Indeed, bounded convergence is merely pointwise convergence with the extra hypothesis that all functions be bounded by the same positive function.

Let us give some examples that distinguish between the notions of convergence we have.

3.5.10 Examples (Pointwise, bounded, and uniform convergence)

1. The sequence of functions in Example 3.5.3 converges pointwise, boundedly, but not uniformly.
2. The sequence of functions in Example 3.5.6 converges pointwise, boundedly, but not uniformly.
3. Consider now a new sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ defined on $I = [0, 1]$ by

$$f_j(x) = \begin{cases} 2j^2x, & x \in [0, \frac{1}{2j}], \\ -2j^2x + 2j, & x \in (\frac{1}{2j}, \frac{1}{j}], \\ 0, & \text{otherwise.} \end{cases}$$

A few members of the sequence are shown in Figure 3.13. This sequence

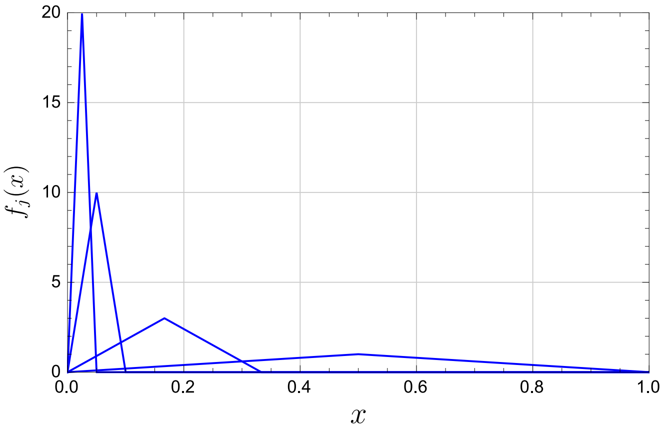


Figure 3.13 A sequence converging pointwise but not boundedly (shown are f_j, $j \in \{1, 5, 10, 20\}$)

converges pointwise to the zero function. Moreover, one can easily check that the convergence is dominated by the function $g\colon [0,1] \to \mathbb{R}$ defined by

$$g(x) = \begin{cases} \frac{1}{x}, & x \in (0,1], \\ 1, & x = 0. \end{cases}$$

The sequence converges neither boundedly nor uniformly.

4. On $I = \mathbb{R}$ consider the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ defined by $f_j(x) = x^2 + \frac{1}{j}$. This sequence clearly converges uniformly to $f\colon x \mapsto x^2$. However, it does not converge boundedly. Of course, the reason is simply that f is itself not bounded. We shall see that uniform convergence to a bounded function implies bounded convergence, in a certain sense. •

We have the following relationship between uniform and bounded convergence.

3.5.11 Proposition (Relationship between uniform and bounded convergence) *If a sequence* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *defined on an interval* I *converges uniformly to a bounded function* f, *then there exists* $\mathrm{N} \in \mathbb{Z}_{>0}$ *such that the sequence* $(\mathrm{f_{N+j}})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges boundedly to* f.

Proof Let $M \in \mathbb{R}_{>0}$ have the property that $|f(x)| < \frac{M}{2}$ for each $x \in I$. Since $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges uniformly to f there exists $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \frac{M}{2}$ for all $x \in I$ and for $j > N$. It then follows that

$$|f_j(x)| \leq |f(x) - f_j(x)| + |f(x)| < M$$

provided that $j > N$. From this the result follows since pointwise convergence of $(f_j)_{j\in\mathbb{Z}_{>0}}$ to f implies pointwise convergence of $(f_{N+j})_{j\in\mathbb{Z}_{>0}}$ to f. ■

3.5.4 Series of ℝ-valued functions

In the previous sections we considered the general matter of sequences of functions. Of course, this discussion carries over to ***series*** of functions, by which we mean expressions of the form $S(x) = \sum_{j=1}^{\infty} f_j(x)$. This is done in the usual manner by considering the partial sums. Let us do this formally.

3.5.12 Definition (Convergence of series) Let $I \subseteq \mathbb{R}$ be an interval and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of ℝ-valued functions on I. Let $F(x) = \sum_{j=1}^{\infty} f_j(x)$ be a series. The corresponding sequence of ***partial sums*** is the sequence $(F_k)_{k\in\mathbb{Z}_{>0}}$ of ℝ-valued functions on I defined by

$$S_k(x) = \sum_{j=1}^{k} f_j(x).$$

Let $g\colon I \to \mathbb{R}_{>0}$. The series:

(i) ***converges pointwise*** if the sequence of partial sums converges pointwise;

(ii) ***converges uniformly*** if the sequence of partial sums converges uniformly;

(iii) ***converges dominated by* g** if the sequence of partial sums converges dominated by g;

(iv) ***converges boundedly*** if the sequence of partial sums converges boundedly. •

A fairly simple extension of pointwise convergence of series is the following notion which is unique to series (as opposed to sequences).

3.5.13 Definition (Absolute convergence of series) Let $I \subseteq \mathbb{R}$ be an interval and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of $\mathbb{R}$-valued functions on I. The sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ ***converges absolutely*** if, for each $x \in I$ and for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $||f_j(x)| - |f_k(x)|| < \epsilon$ provided that $j, k \geq N$. •

Thus an absolutely convergent sequence is one where, for each $x \in I$, the sequence $(|f_j(x)|)_{j\in\mathbb{Z}_{>0}}$ is Cauchy, and hence convergent. In other words, for each $x \in I$, the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ is absolutely convergent. It is clear, then, that an absolutely convergent sequence of functions is pointwise convergent. Let us give some examples that illustrate the difference between pointwise and absolute convergence.

3.5.14 Examples (Absolute convergence)

1. The sequence of functions of Example 3.5.3 converges absolutely since the functions all take positive values.
2. For $j \in \mathbb{Z}_{>0}$, define $f_j\colon [0,1] \to \mathbb{R}$ by $f_j(x) = \frac{(-1)^{j+1}x}{j}$. Then, by Example 2.4.2–3, the series $S(x) = \sum_{j=1}^{\infty} f_j(x)$ is absolutely convergent if and only $x = 0$. But in Example 2.4.2–3 we showed that the series is pointwise convergent. •

3.5.5 Some results on uniform convergence of series

At various times in our development, we will find it advantageous to be able to refer to various standard results on uniform convergence, and we state these here.

Let us first recall the Weierstrass M-test.

3.5.15 Theorem (Weierstrass M-test) *If* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a sequence of* $\mathbb{R}$*-valued functions defined on an interval* $\mathrm{I} \subseteq \mathbb{R}$ *and if there exists a sequence of positive constants* $(\mathrm{M_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *such that*

(i) $|\mathrm{f_j}(\mathrm{x})| \leq \mathrm{M_j}$ *for all* $\mathrm{x} \in \mathrm{I}$ *and for all* $\mathrm{j} \in \mathbb{Z}_{>0}$ *and*

(ii) $\sum_{\mathrm{j}=1}^{\infty} \mathrm{M_j} < \infty$,

then the series $\sum_{\mathrm{j}=1}^{\infty} \mathrm{f_j}$ *converges uniformly and absolutely.*

Proof For $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that, if $l \geq N$, we have

$$|M_l + \dots + M_{l+k}| < \epsilon$$

for every $k \in \mathbb{Z}_{>0}$. Therefore, by the triangle inequality,

$$\left|\sum_{j=l}^{l+k} f_j(x)\right| \leq \sum_{j=l}^{l+k} |f_j(x)| \leq \sum_{j=l}^{l+k} M_j.$$

This shows that, for every $\epsilon \in \mathbb{R}_{>0}$, the tail of the series $\sum_{j=1}^{\infty} f_j$ can be made smaller than ϵ, and uniformly in x. This implies uniform and absolute convergence. ■

Next we present Abel's test.

3.5.16 Theorem (Abel's test) *Let $(g_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of $\mathbb{R}$-valued functions on an interval $I \subseteq \mathbb{R}$ for which $g_{j+1}(x) \leq g_j(x)$ for all $j \in \mathbb{Z}_{>0}$ and $x \in I$. Also suppose that there exists $M \in \mathbb{R}_{>0}$ such that $g_j(x) \leq M$ for all $x \in I$ and $j \in \mathbb{Z}_{>0}$. Then, if the series $\sum_{j=1}^{\infty} f_j$ converges uniformly on I, then so too does the series $\sum_{j=1}^{\infty} g_j f_j$.*

Proof Denote

$$F_k(x) = \sum_{j=1}^{k} f_j(x), \quad G_k(x) = \sum_{j=1}^{k} g_j(x) f_j(x)$$

as the partial sums. Using Abel's partial summation formula (Proposition 2.4.16), for $0 < k < l$ we write

$$G_l(x) - G_k(x) = (F_l(x) - F_k(x))G_1(x) + \sum_{j=k+1}^{l} (F_l(x) - F_j(x))(g_{j+1}(x) - g_j(x)).$$

An application of the triangle inequality gives

$$|G_l(x) - G_k(x)| = \big|(F_l(x) - F_k(x))\big||G_1(x)| + \sum_{j=k+1}^{l} \big|(F_l(x) - F_j(x))\big|(g_{j+1}(x) - g_j(x)),$$

since $|g_{j+1}(x) - g_j(x)| = g_{j+1}(x) - g_j(x)$. Now, given $\epsilon \in \mathbb{R}_{>0}$, let $N \in \mathbb{Z}_{>0}$ have the property that

$$\big|F_l(x) - F_k(x)\big| \leq \frac{\epsilon}{3M}$$

for all $k, l \geq N$. Then we have

$$\begin{aligned} |G_l(x) - G_k(x)| &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3M} \sum_{j=k+1}^{l} (g_{j+1}(x) - g_j(x)) \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3M}(g_{k+1}(x) - g_{l+1}(x)) \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3M}(|g_{k+1}(x)| + |g_{l+1}(x)|) \leq \epsilon. \end{aligned}$$

Thus the sequence $(G_j)_{j\in\mathbb{Z}_{>0}}$ is uniformly Cauchy, and hence uniformly convergent. ■

The final result on general uniform convergence we present is the Dirichlet test.[12]

3.5.17 Theorem (Dirichlet's test) *Let $(f_j)_{j\in\mathbb{Z}_{>0}}$ and $(g_j)_{j\in\mathbb{Z}_{>0}}$ be sequences of $\mathbb{R}$-valued functions on an interval I and satisfying the following conditions:*

(i) there exists $M \in \mathbb{R}_{>0}$ such that the partial sums

$$F_k(x) = \sum_{j=1}^{k} f_j(x)$$

satisfy $|F_k(x)| \leq M$ for all $k \in \mathbb{Z}_{>0}$ and $x \in I$;

[12] Johann Peter Gustav Lejeune Dirichlet 1805–1859 was born in what is now Germany. His mathematical work was primarily in the areas of analysis, number theory and mechanics. For the purposes of these volumes, Dirichlet was gave the first rigorous convergence proof for the trigonometric series of Fourier. These and related results are presented in Section 12.2.

(ii) $\mathrm{g_j(x)} \geq 0$ *for all* $\mathrm{j} \in \mathbb{Z}_{>0}$ *and* $\mathrm{x} \in \mathrm{I}$*;*

(iii) $\mathrm{g_{j+1}(x)} \leq \mathrm{g_j(x)}$ *for all* $\mathrm{j} \in \mathbb{Z}_{>0}$ *and* $\mathrm{x} \in \mathrm{I}$*;*

(iv) the sequence $(\mathrm{g_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges uniformly to the zero function.*

Then the series $\sum_{\mathrm{j}=1}^{\infty} \mathrm{f_j g_j}$ *converges uniformly on* I*.*

Proof We denote

$$F_k(x) = \sum_{j=1}^{k} f_j(x), \quad G_k(x) = \sum_{j=1}^{k} f_j(x)g_j(x).$$

We use again the Abel partial summation formula, Proposition 2.4.16, to write***missing stuff***

$$G_l(x) - G_k(x) = F_l(x)g_{l+1}(x) - F_k(x)g_{k+1}(x) - \sum_{j=k+1}^{l} F_j(x)(g_{l+1}(x) - g_l(x)).$$

Now we compute

$$\begin{aligned}|G_l(x) - G_k(x)| &\leq M(g_{l+1}(x) + g_{k+1}(x)) + M\sum_{j=k+1}^{l}(g_j(x) - g_{j+1}(x))\\ &= 2Mg_{k+1}(x).\end{aligned}$$

Now, for $\epsilon \in \mathbb{R}_{>0}$, if one chooses $N \in \mathbb{Z}_{>0}$ such that $g_k(x) \leq \frac{\epsilon}{2M}$ for all $x \in I$ and $k \geq N$, then it follows that $|G_l(x) - G_k(x)| \leq \epsilon$ for $k, l \geq N$ and for all $x \in I$. From this we deduce that the sequence of partial sums $(G_j)_{j\in\mathbb{Z}_{>0}}$ is uniformly Cauchy, and hence uniformly convergent. ■

3.5.6 The Weierstrass Approximation Theorem

In this section we prove an important result in analysis. The theorem is one on approximating continuous functions with a certain class of easily understood functions. The idea, then, is that if one say something about the class of easily understood functions, it may be readily also ascribed to continuous functions. Let us first describe the class of functions we wish to use to approximate continuous functions.

3.5.18 Definition (Polynomial functions) A function $P\colon \mathbb{R} \to \mathbb{R}$ is a ***polynomial function*** if

$$P(x) = a_k x^k + \cdots + a_1 x + a_0$$

for some $a_0, a_1, \dots, a_k \in \mathbb{R}$. The ***degree*** of the polynomial function P is the largest $j \in \{0, 1, \dots, k\}$ for which $a_j \neq 0$. •

We shall have a great deal to say about polynomials in an algebraic setting in Section **??**. Here we will only think about the most elementary features of polynomials.

Our constructions are based on a special sort of polynomial. We recall the notation

$$\binom{m}{k} \triangleq \frac{m!}{k!(m-k)!}$$

which are the ***binomial coefficients***.

3.5.19 Definition (Bernstein polynomial, Bernstein approximation) For $m \in \mathbb{Z}_{\geq 0}$ and $k \in \{0, 1, \dots, m\}$ the polynomial function

$$P_k^m(x) = \binom{m}{k} x^k (1-x)^{m-k}$$

is a ***Bernstein polynomial***. For a continuous function $f\colon [a,b] \to \mathbb{R}$ the **m*th Bernstein approximation*** of f is the function $B_m^{[a,b]} f\colon [a,b] \to \mathbb{R}$ defined by

$$B_m^{[a,b]} f(x) = \sum_{k=0}^{m} f(a + \tfrac{k}{m}(b-a)) P_k^m(\tfrac{x-a}{b-a}). \qquad \bullet$$

In Figure 3.14 we depict some of the Bernstein polynomials. The way to imagine

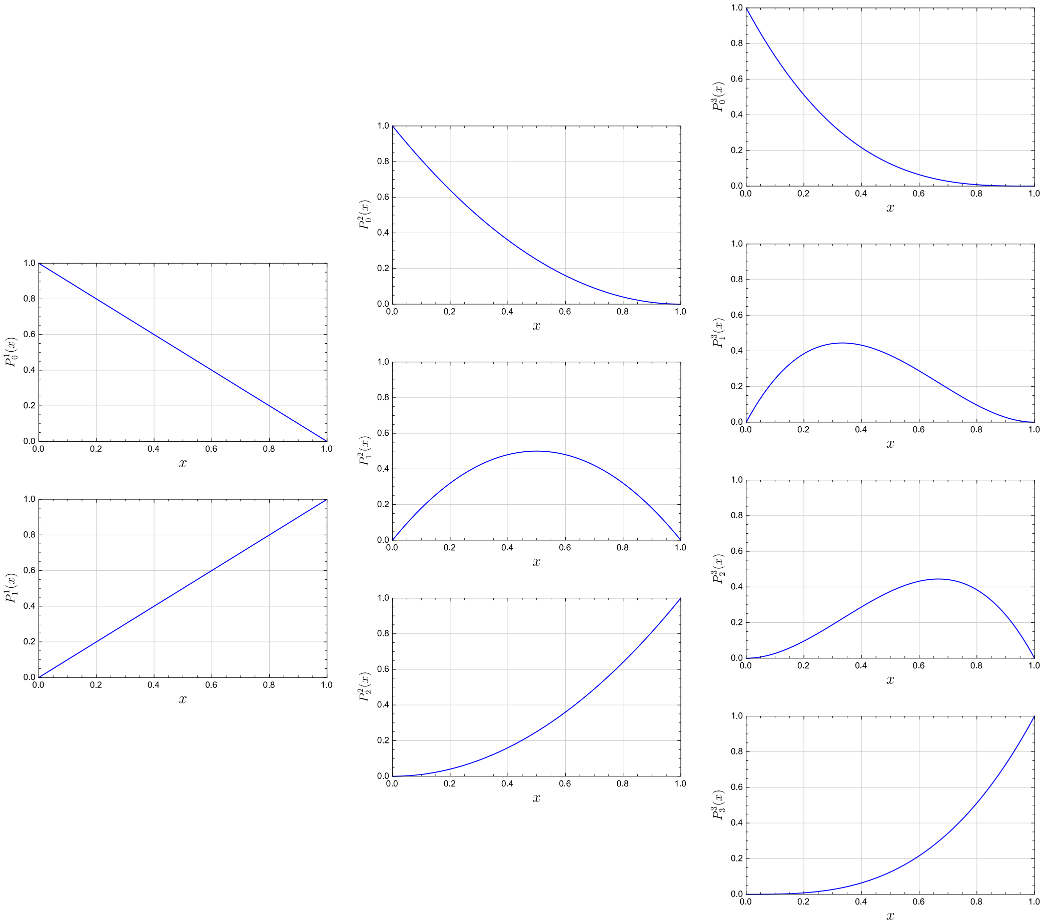


Figure 3.14 The Bernstein polynomials P_0^1 and P_1^1 (left), P_0^2, P_1^2, and P_2^2 (middle), and P_0^3, P_1^3, P_2^3, and P_3^3 (right)

the point of these functions is as follows. The polynomial P_k^m on the interval $[0,1]$ has a single maximum at $\frac{k}{m}$. By letting m vary over $\mathbb{Z}_{\geq 0}$ and letting $k \in \{0,1,\dots,m\}$, the points of the form $\frac{k}{m}$ will get arbitrarily close to any point in $[0,1]$. The function $f(\frac{k}{m})P_k^m$ thus has a maximum at $\frac{k}{m}$ and the behaviour of f away from $\frac{k}{m}$ is thus (sort of) attenuated. In fact, for large m the behaviour of the function P_k^m becomes increasingly "focussed" at $\frac{k}{m}$. Thus, as m gets large, the function $f(\frac{k}{m})P_k^m$ starts looking like the function taking the value $f(\frac{k}{m})$ at $\frac{k}{m}$ and zero elsewhere. Now, using the identity

$$\sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^m = 1 \tag{3.16}$$

which can be derived using the Binomial Theorem (see Exercise 2.2.1), this means that for large m, $B_m^{[0,1]} f(\frac{k}{m})$ approaches the value $f(\frac{k}{m})$. This is the idea of the Bernstein approximation.

That being said, let us prove some basic facts about Bernstein approximations.

3.5.20 Lemma (Properties of Bernstein approximations) *For continuous functions* $f, g\colon [a,b] \to \mathbb{R}$, *for* $\alpha \in \mathbb{R}$, *and for* $m \in \mathbb{Z}_{\geq 0}$, *the following statements hold:*

(i) $B_m^{[a,b]}(f+g) = B_m^{[a,b]}f + B_m^{[a,b]}g$;

(ii) $B_m^{[a,b]}(\alpha f) = \alpha B_m^{[a,b]}f$;

(iii) $B_m^{[a,b]}f(x) \geq 0$ *for all* $x \in [a,b]$ *if* $f(x) \geq 0$ *for all* $x \in [a,b]$;

(iv) $B_m^{[a,b]}f(x) \leq B_m^{[a,b]}g(x)$ *for all* $x \in [a,b]$ *if* $f(x) \leq g(x)$ *for all* $x \in [a,b]$;

(v) $|B_m^{[a,b]}f(x)| \leq B_m^{[a,b]}g(x)$ *for all* $x \in [a,b]$ *if* $|f(x)| \leq g(x)$ *for all* $x \in [a,b]$;

(vi) *for* $k, m \in \mathbb{Z}_{\geq 0}$ *we have*

$$(B_{m+k}^{[a,b]})^{(k)}(x) = \frac{(m+k)!}{m!} \frac{1}{(b-a)^k} \sum_{j=0}^{m} \Delta_h^k f(a + \tfrac{j}{k+m}(b-a)) P_j^m(\tfrac{x-a}{b-a}),$$

where $h = \frac{1}{k+m}$ *and where* $\Delta_h^k f\colon [a,b] \to \mathbb{R}$ *is defined by*

$$\Delta_h^k f(x) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} f(x + jh)$$

(vii)

(viii) *if we define* $f_0, f_1, f_2\colon [0,1] \to \mathbb{R}$ *by*

$$f_0(x) = 1, \quad f_1(x) = x, \quad f_2(x) = x^2, \qquad x \in [0,1],$$

then

$$B_m^{[0,1]}f_0(x) = 1, \quad B_m^{[0,1]}f_1(x) = x, \quad B_m^{[0,1]}f_2(x) = x^2 + \tfrac{1}{m}(x - x^2)$$

for $x \in [0,1]$ *and* $m \in \mathbb{Z}_{\geq 0}$.

Proof Let $\hat{f}\colon [0,1] \to \mathbb{R}$ be defined by $\hat{f}(y) = f(a + \frac{y}{(}b - a))$. One can verify that if the lemma holds for $\hat{f}$ then it immediately follows for f, and so without loss of generality we suppose that $[a,b] = [0,1]$. We also abbreviate $B_m^{[0,1]} = B_m$.

(i)–(iv) These assertions follow directly from the definition of the Bernstein approximations.

(v) If $|f(x)| \le g(x)$ for all $x \in [0,1]$ then

$$\begin{aligned} & -f(x) \le g(x) \le f(x), \qquad x \in [0,1] \\ \implies \quad & -B_m f(x) \le B_m g(x) \le B_m f(x), \qquad x \in [0,1], \end{aligned}$$

using the fourth assertion.

(vi) Note that

$$B_{m+k}(x) = \sum_{j=0}^{m+k} f(\tfrac{j}{m+k}) \binom{m+k}{j} x^j (1-x)^{m+k-j}.$$

Let $g_j(x) = x^j$ and $h_j(x) = (1-x)^{m+k-j}$ and compute

$$g_j^{(r)}(x) = \begin{cases} \frac{j!}{(j-r)!} x^{j-r}, & j - r \ge 0, \\ 0, & j - r < 0 \end{cases}$$

and

$$h_j^{(k-r)}(x) = \begin{cases} (-1)^{k-r} \frac{(m+k-j)!}{(m+r-j)!} (1-x)^{m+r-j}, & j - r \le m, \\ 0, & j - r > m. \end{cases}$$

By Proposition 3.2.11,

$$(g_j h_j)^{(k)}(x) = \sum_{r=0}^{k} \binom{k}{r} g_j^{(r)}(x) h_j^{(k-r)}(x).$$

Also note that

$$\begin{aligned} \binom{m+k}{j} \frac{j!}{(j-r)!} \frac{(m+k-j)!}{(m+r-j)!} &= \frac{(m+k)!}{j!(m+k-j)!} \frac{j!}{(j-r)!} \frac{(m+k-j)!}{(m+r-j)!} \\ &= \frac{(m+k)!}{m!} \frac{m!}{(m-(j-r))!(j-r)!} = \frac{(m+k)!}{m!} \binom{m}{j-r}. \end{aligned}$$

Putting this all together we have

$$\begin{aligned} B_{m+k}^{(k)}(x) &= \sum_{j=0}^{m+k} \sum_{r=0}^{k} f(\tfrac{j}{m+k}) \binom{m+k}{j} \binom{k}{r} g_j^{(r)}(x) h_j^{(k-r)}(x) \\ &= \sum_{r=0}^{k} \sum_{l=-r}^{m+k-r} f(\tfrac{l+r}{m+k}) \binom{m+k}{l+r} \binom{k}{r} g_{l+r}^{(r)}(x) h_{l+r}^{(k-r)}(x) \\ &= \sum_{r=0}^{k} \sum_{l=0}^{m} (-1)^{k-r} \binom{k}{r} f(\tfrac{l+r}{m+k}) \binom{m}{l} x^l (1-x)^{n-l}, \end{aligned}$$

where we make the change of index $(l, r) = (j - r, r)$ in the second step and note that the derivatives of g_{l+r} and h_{l+r} vanish when $l < 0$ and $l > m$. Let $h = \frac{1}{m+k}$. Since

$$\Delta_h^k f(\tfrac{j}{m+k}) = \sum_{r=0}^{k} (-1)^{k-r} \binom{k}{r} f(\tfrac{j+r}{m+k})$$

this part of the result follows.

(vii)

(viii) It follows from (3.16) that $B_m f_0(x) = 1$ for every $x \in [0, 1]$. We also compute

$$\begin{aligned} B_m f_0(x) &= \sum_{k=0}^{m} \frac{k}{m} \frac{m!}{m!(m-k)!} x^k (1-x)^{m-k} \\ &= x \sum_{k=0}^{m-1} \frac{(m-1)!}{(k-1)!((m-1)-(k-1))!} x^k (1-x)^{m-1-k} \\ &= x(x + (1-x))^{m-1} = x, \end{aligned}$$

where we use the Binomial Theorem. To compute $B_m f_2$ we first compute

$$\begin{aligned} \frac{k^2}{m^2} \frac{m!}{k!(m-k)!} &= \frac{(k-1)+1}{m} \frac{(m-1)!}{(k-1)!(m-k)!} \\ &= \frac{(k-1)(n-1)}{n(n-1)} \frac{(m-1)!}{(k-1)!(m-k)!} + \frac{1}{m} \frac{(m-1)!}{(k-1)!(m-k)!} \\ &= \frac{m-1}{m} \binom{n-2}{k-2} + \frac{1}{m} \binom{n-1}{k-1}, \end{aligned}$$

where we adopt the convention that $\binom{j}{l} = 0$ if either j or l are zero. We now compute

$$\begin{aligned} B_m f_2(x) &= \sum_{k=0}^{m} \frac{k^2}{m^2} \binom{m}{k} x^k (1-x)^{m-k} \\ &= \frac{m-1}{m} \sum_{k=2}^{m} \binom{m-2}{k-2} x^k (1-x)^{m-k} + \frac{1}{m} \sum_{k=1}^{m} \binom{m-1}{k-1} x^k (1-x)^{m-k} \\ &= \frac{m-1}{m} x^2 (x + (1-x))^{m-2} + \frac{1}{m} x (x + (1-x))^{m-1} = \frac{m-1}{m} x^2 + \frac{1}{m} x, \end{aligned}$$

as desired. ■

Now, heuristics aside, we state the main result in this section, a consequence of which is that every continuously function on a compact interval can be approximated arbitrarily well (in the sense that the maximum difference can be made as small as desired) by a polynomial function.

3.5.21 Theorem (Weierstrass Approximation Theorem) *Consider a compact interval* $[a, b] \subseteq \mathbb{R}$ *and let* $f\colon [a, b] \to \mathbb{R}$ *be continuous. Then the sequence* $(B_m^{[a,b]} f)_{m \in \mathbb{Z}_{>0}}$ *converges uniformly to* f *on* $[a, b]$.

Proof It is evident (why?) that we can take $[a,b] = [0,1]$ and then let us denote $B_m f = B_m^{[0,1]} f$ for simplicity.

Let $\epsilon \in \mathbb{R}_{>0}$. Since f is uniformly continuous by Theorem 3.1.24 there exists $\delta \in \mathbb{R}_{>0}$ such that $|f(x) - f(y)| \leq \frac{\epsilon}{2}$ whenever $|x - y| \leq \delta$. Let

$$M = \sup\{|f(x)| \mid x \in [0,1]\},$$

noting that $M < \infty$ by Theorem 3.1.23. Note then that if $|x - y| \leq \delta$ then

$$|f(x) - f(y)| \leq \tfrac{\epsilon}{2} \leq \tfrac{\epsilon}{2} + \tfrac{2M}{\delta^2}(x-y)^2.$$

If $|x - y| > \delta$ then

$$|f(x) - f(y)| \leq 2M \leq 2M\Big(\tfrac{x-y}{\delta}\Big)^2 \leq \tfrac{\epsilon}{2} + \tfrac{2M}{\delta^2}(x-y)^2.$$

That is to say, for every $x, y \in [0,1]$,

$$|f(x) - f(y)| \leq \tfrac{\epsilon}{2} + \tfrac{2M}{\delta^2}(x-y)^2. \tag{3.17}$$

Now, fix $x_0 \in [0,1]$ and compute, using the lemma above (along with the notation f_0, f_1, and f_2 introduced in the lemma) and (3.17),

$$\begin{aligned}|B_m f(x) - f(x_0)| = |B_m(f - f(x_0)f_0)(x)| &\leq B_m\Big(\tfrac{\epsilon}{2} f_0 + \tfrac{2M}{\delta^2}(f_1 - x_0 f_0)^2\Big)(x)\\ &= \tfrac{\epsilon}{2} + \tfrac{2M}{\delta^2}(x^2 + \tfrac{1}{m}(x - x^2) - 2x_0 x + x_0^2)\\ &= \tfrac{\epsilon}{2} + \tfrac{2M}{\delta^2}(x - x_0)^2 + \tfrac{2M}{m\delta^2}(x - x^2),\end{aligned}$$

this holding for every $m \in \mathbb{Z}_{\geq 0}$. Now evaluate at $x = x_0$ to get

$$|B_m f(x_0) - f(x_0)| \leq \tfrac{\epsilon}{2} + \tfrac{2M}{m\delta^2}(x_0 - x_0^2) \leq \tfrac{\epsilon}{2} + \tfrac{M}{2m\delta^2},$$

using the fact that $x_0 - x_0^2 \leq \frac{1}{4}$ for $x_0 \in [0,1]$. Therefore, if $N \in \mathbb{Z}_{>0}$ is sufficiently large that $\frac{M}{2m\delta^2} < \frac{\epsilon}{2}$ for $m \geq N$ we have

$$|B_m f(x_0) - f(x_0)| < \epsilon,$$

and this holds for every $x_0 \in [0,1]$, giving us the desired uniform convergence. ■

For fun, let us illustrate the Bernstein approximations in an example.

3.5.22 Example (Bernstein approximation) Let us consider $f\colon [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x, & x \in [0, \frac{1}{2}],\\ 1 - x, & x \in (\frac{1}{2}, 1].\end{cases}$$

In Figure 3.15 we show some Bernstein approximations to f. Note that the convergence is rather poor. One might wish to contrast the 100th approximation in Figure 3.15 with the 10 approximation of the same function using Fourier series depicted in Figure 12.11. (If you have no clue what a Fourier series is, that is fine. We will get there in time.) •

We shall revisit the Weierstrass Approximation Theorem in Sections **??** and ***missing stuff***.

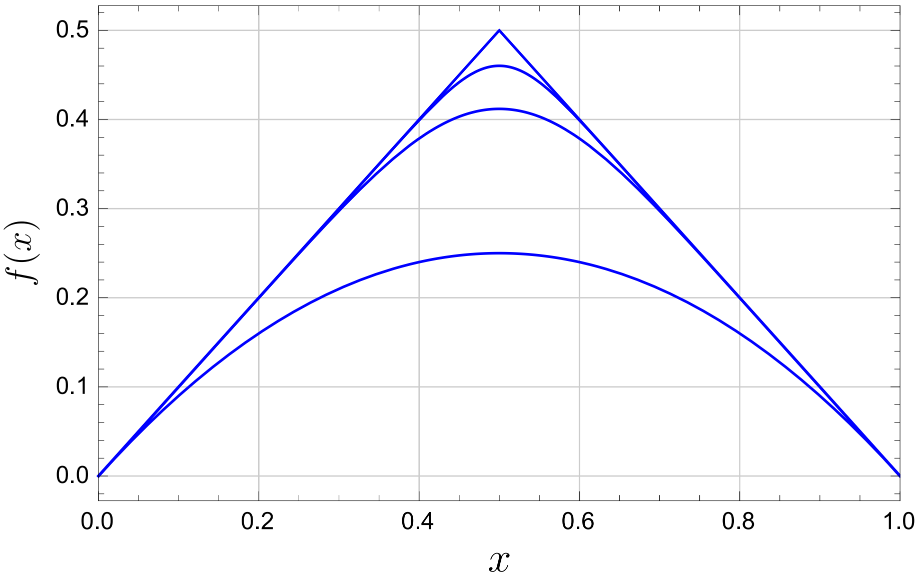


Figure 3.15 Bernstein approximations for $m \in \{2, 50, 100\}$

3.5.7 Swapping limits with other operations

In this section we give some basic result concerning the swapping of various function operations with limits. The first result we consider pertains to integration. When we consider Lebesgue integration in Chapter 5 we shall see that there are more powerful limit theorems available. Indeed, the *raison d'etre* for the Lebesgue integral is just these limit theorems, as these are not true for the Riemann integral. However, for the moment these theorems have value in that they apply in at least some cases, and indicate what *is* true for the Riemann integral.

3.5.23 Theorem (Uniform limits commute with Riemann integration) *Let* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *be a compact interval and let* $(\mathrm{f}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be a sequence of continuous* $\mathbb{R}$*-valued functions defined on* $[\mathrm{a}, \mathrm{b}]$ *that converge uniformly to* f. *Then*

$$\lim_{\mathrm{j}\to\infty} \int_\mathrm{a}^\mathrm{b} \mathrm{f}_\mathrm{j}(\mathrm{x})\,\mathrm{dx} = \int_\mathrm{a}^\mathrm{b} \mathrm{f}(\mathrm{x})\,\mathrm{dx}.$$

Proof As the functions $(f_j)_{j\in\mathbb{Z}_{>0}}$ are continuous and the convergence to f is uniform, f must be continuous by Theorem 3.5.8. Since the interval $[a, b]$ is compact, the functions f and f_j, $j \in \mathbb{Z}_{>0}$, are also bounded. Therefore, by part Proposition 3.4.25,***missing stuff***

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \le M(b-a)$$

where $M = \sup\{|f(x)| \mid x \in [a, b]\}$. Let $\epsilon \in \mathbb{R}_{>0}$ and select $N \in \mathbb{Z}_{>0}$ such that $|f_j(x) - f(x)| <$

$\frac{\epsilon}{b-a}$ for all $x \in [a,b]$, provided that $j \geq N$. Then

$$\left|\int_a^b f_j(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x\right| = \left|\int_a^b (f_j(x) - f(x))\,\mathrm{d}x\right| \\ \leq \frac{\epsilon}{b-a}(b-a) = \epsilon.$$

This is the desired result. ■

Next we state a result that tells us when we may switch limits and differentiation.

3.5.24 Theorem (Uniform limits commute with differentiation) *Let* $I = [a,b]$ *be a compact interval and let* $(f_j)_{j\in\mathbb{Z}_{>0}}$ *be a sequence continuously differentiable* ℝ*-valued functions on* $[a,b]$*, and suppose that the sequence converges pointwise to* f*. Also suppose that the sequence* $(f_j')_{j\in\mathbb{Z}_{>0}}$ *of derivatives converges uniformly to* g*. Then* f *is differentiable and* $f' = g$*.*

Proof Our hypotheses ensure that we may write, for each $j \in \mathbb{Z}_{>0}$,

$$f_j(x) = f_j(a) + \int_a^x f_j'(\xi)\,\mathrm{d}\xi.$$

for each $x \in [a,b]$. By Theorem 3.5.23, we may interchange the limit as $j \to \infty$ with the integral, and so we get

$$f(t) = f(a) + \int_a^x g(\xi)\,\mathrm{d}\xi.$$

Since g is continuous, being the uniform limit of continuous functions (by Theorem 3.5.8), the Fundamental Theorem of Calculus ensures that $f' = g$. ■

The next result in this section has a somewhat different character than the rest. It actually says that it is possible to differentiate a sequence of monotonically increasing functions term-by-term, except on a set of measure zero. The interesting thing here is that only pointwise convergence is needed.

3.5.25 Theorem (Termwise differentiation of sequences of monotonic functions is a.e. valid) *Let* $I = [a,b]$ *be a compact interval, let* $(f_j)_{j\in\mathbb{Z}_{>0}}$ *be a sequence of monotonically increasing functions such that the series* $S = \sum_{j=1}^{\infty} f_j(x)$ *converges pointwise to a function* f*. Then there exists a set* $Z \subseteq I$ *such that*

(i) Z *has measure zero and*

(ii) $f'(x) = \sum_{j=1}^{\infty} f_j'(x)$ *for all* $x \in I \setminus Z$*.*

Proof Note that the limit function f is monotonically increasing. Denote by $Z_1 \subseteq [a,b]$ the set of points for which all of the functions f and f_j, $j \in \mathbb{Z}_{>0}$, do not possess derivatives. Note that by Theorem 3.2.26 it follows that Z_1 is a countable union of sets of measure zero. Therefore, by Exercise 2.5.9, Z_1 has measure zero. Now let $x \in I \setminus Z_1$ and let $\epsilon \in \mathbb{R}_{>0}$ be sufficiently small that $x + \epsilon \in [a,b]$. Then

$$\frac{f(x+\epsilon) - f(x)}{\epsilon} = \sum_{j=1}^{\infty} \frac{f_j(x+\epsilon) - f_j(x)}{\epsilon}.$$

Since $f_j(x+\epsilon) - f_j(x) \geq 0$, for any $k \in \mathbb{Z}_{>0}$ we have

$$\frac{f(x+\epsilon) - f(x)}{\epsilon} \geq \sum_{j=1}^{k} \frac{f_j(x+\epsilon) - f_j(x)}{\epsilon},$$

which then gives

$$f'(x) \geq \sum_{j=1}^{k} f_j'(x).$$

The sequence of partial sums for the series $\sum_{j=1}^{\infty} f_j'(x)$ is therefore bounded above. Moreover, by Theorem 3.2.26, it is increasing. Therefore, by Theorem 2.3.8 the series $\sum_{j=1}^{\infty} f_j'(x)$ converges for every $x \in I \setminus Z_1$.

Let us now suppose that $f(a) = 0$ and $f_j(a) = 0$, $j \in \mathbb{Z}_{>0}$. This can be done without loss of generality by replacing f with $f - f(a)$ and f_j with $f_j - f_j(a)$, $j \in \mathbb{Z}_{>0}$. With this assumption, for each $x \in [a,b]$ and $k \in \mathbb{Z}_{>0}$, we have $f(x) - S_k(x) \geq 0$ where $(S_k)_{k\in\mathbb{Z}_{>0}}$ is the sequence of partial sums for S. Choose a subsequence $(S_{k_l})_{l\in\mathbb{Z}_{>0}}$ of $(S_k)_{k\in\mathbb{Z}_{>0}}$ having the property that $0 \leq f(b) - S_{k_l}(b) \leq 2^{-l}$, this being possible since the sequence $(S_k(b))_{k\in\mathbb{Z}_{>0}}$ converges to $f(b)$. Note that

$$f(x) - S_{k_l}(x) = \sum_{j=k_l+1}^{\infty} f_j(x),$$

meaning that $f - S_{k_l}$ is a monotonically increasing function. Therefore, $0 \leq f(x) - S_{k_l}(x) \leq 2^{-l}$ for all $x \in [a,b]$. This shows that the series $\sum_{l=1}^{\infty}(f(x) - S_{k_l}(x))$ is a pointwise convergent sequence of monotonically increasing functions. Let g denote the limit function, and let $Z_2 \subseteq [a,b]$ be the set of points where all of the functions g and $f - S_{k_l}$, $l \in \mathbb{Z}_{>0}$, do not possess derivatives, noting that this set is, in the same manner as was Z_1, a set of measure zero. The argument above applies again to show that, for $x \in I \setminus Z_2$, the series $\sum_{l=1}^{\infty}(f'(x) - S_{k_l}'(x))$ converges. Thus, for $x \in I \setminus Z_2$, it follows that $\lim_{l\to\infty}(f'(x) - S_{k_l}'(x)) = 0$. Now, for $x \in I \setminus Z_1$, we know that $(S_k'(x))_{k\in\mathbb{Z}_{>0}}$ is a monotonically increasing sequence. Therefore, for $x \in I \setminus (Z_1 \cup Z_2)$, the sequence $(f'(x) - S_k'(x))_{k\in\mathbb{Z}_{>0}}$ must converge to zero. This gives the result by taking $Z = Z_1 \cup Z_2$. ■

As a final result, we indicate how convexity interacts with pointwise limits.

3.5.26 Theorem (The pointwise limit of convex functions is convex) *If* $I \subseteq \mathbb{R}$ *is convex and if* $(f_j)_{j\in\mathbb{Z}_{>0}}$ *is a sequence of convex functions converging pointwise to* $f\colon I \to \mathbb{R}$, *then* f *is convex.*

Proof Let $x_1, x_2 \in I$ and let $s \in [0,1]$. Then

$$\begin{aligned} f((1-s)x_1 + sx_2) = \lim_{j\to\infty} f_j((1-s)x_1 + sx_2) &\leq \lim_{j\to\infty}((1-s)f_j(x_1) + sf_j(x_2)) \\ &= (1-s)\lim_{j\to\infty} f_j(x_1) + s\lim_{j\to\infty} f_j(x_2) \\ = (1-s)f(x_1) + sf(x_2), \end{aligned}$$

where we have used Proposition 2.3.23. ■

3.5.8 Notes

There are many proofs available of the Weierstrass Approximation Theorem, and the rather explicit proof we give is due to **SNB:12**.

Exercises

3.5.1 Consider the sequence of functions $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ defined on the interval $[0,1]$ by $f_j(x) = x^{1/2^j}$. Thus

$$f_1(x) = \sqrt{x},\ \ f_2(x) = \sqrt{f_1(x)} = \sqrt{\sqrt{x}},\ \ \dots,\ \ f_j(x) = \sqrt{f_{j-1}(x)} = x^{1/2^j},\dots$$

(a) Sketch the graph of f_j for $j \in \{1,2,3\}$.

(b) Does the sequence of functions $(f_j)_{j\in\mathbb{Z}_{>0}}$ converge pointwise? If so, what is the limit function?

(c) Is the convergence of the sequence of functions $(f_j)_{j\in\mathbb{Z}_{>0}}$ uniform?

(d) Is it true that

$$\lim_{j\to\infty}\int_0^1 f_j(x)\,\mathrm{d}x = \int_0^1 \lim_{j\to\infty} f_j(x)\,\mathrm{d}x?$$

3.5.2 In each of the following exercises, you will be given a sequence of functions defined on the interval $[0,1]$. In each case, answer the following questions.

1. Sketch the first few functions in the sequence.
2. Does the sequence converge pointwise? If so, what is the limit function?
3. Does the sequence converge uniformly?

The sequences are as follows:

(a) $(f_j(x) = (x - \frac{1}{j^2})^2)_{j\in\mathbb{Z}_{>0}}$;

(b) $(f_j(x) = x - x^j)_{j\in\mathbb{Z}_{>0}}$.

3.5.3 Let $I \subseteq \mathbb{R}$ be an interval and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of locally bounded functions on I converging pointwise to $f\colon I \to \mathbb{R}$. Show that there exists a function $g\colon I \to \mathbb{R}$ such that $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges dominated by g.

Section 3.6

Some ℝ-valued functions of interest

In this section we present, in a formal way, some of the special functions that will, and indeed already have, come up in these volumes.

Do I need to read this section? It is much more than likely the case that the reader has already encountered the functions we discuss in this section. However, it may be the case that the formal definitions and rigorous presentation of their properties will be new. This section, therefore, fits into the "read for pleasure" category. •

3.6.1 The exponential function

One of the most important functions in mathematics, particularly in applied mathematics, is the exponential function. This importance is nowhere to be found in the following definition, but hopefully at the end of their reading these volumes, the reader will have some appreciation for the exponential function.

3.6.1 Definition (Exponential function) The ***exponential function***, denoted by $\exp\colon \mathbb{R} \to \mathbb{R}$, is given by

$$\exp(x) = \sum_{j=0}^{\infty} \frac{x^j}{j!}. \qquad \bullet$$

In Figure 3.16 we show the graphs of exp and its inverse log that we will be discussing in the next section.

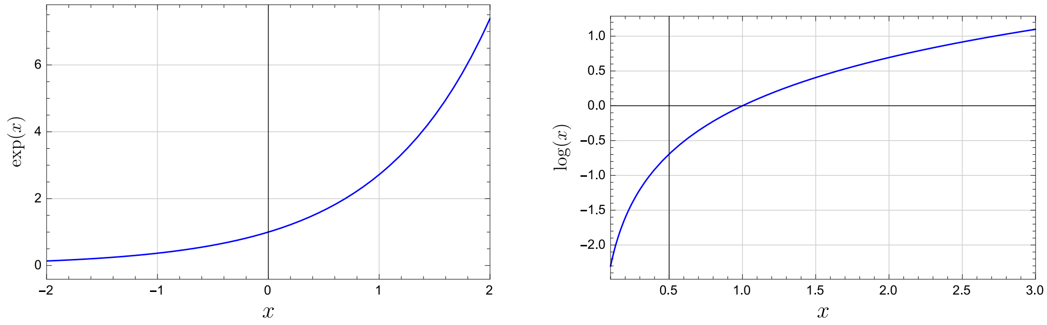


Figure 3.16 The function exp (left) and its inverse log (right)

One can use Theorem **??**, along with Proposition 2.4.15, to easily show that the power series for exp has an infinite radius of convergence, and so indeed defines a function on ℝ. Let us record some of the more immediate and useful properties of exp.

3.6.2 Proposition (Properties of the exponential function) *The exponential function enjoys the following properties:*

(i) $\exp$ *is infinitely differentiable;*

(ii) $\exp$ *is strictly monotonically increasing;*

(iii) $\exp(x) > 0$ *for all* $x \in \mathbb{R}$*;*

(iv) $\lim_{x\to\infty} \exp(x) = \infty$*;*

(v) $\lim_{x\to-\infty} \exp(x) = 0$*;*

(vi) $\exp(x + y) = \exp(x)\exp(y)$ *for all* $x, y \in \mathbb{R}$*;*

(vii) $\exp' = \exp$*;*

(viii) $\lim_{x\to\infty} x^k \exp(-x) = 0$ *for all* $k \in \mathbb{Z}_{>0}$.

Proof (i) This follows from Corollary **??**, along with the fact that the radius of convergence of the power series for exp is infinite.

(vi) Using the Binomial Theorem and Proposition 2.4.30(iv) we compute

$$\begin{aligned}\exp(x)\exp(y) &= \Big(\sum_{j=0}^{\infty} \frac{x^j}{j!}\Big)\Big(\sum_{j=0}^{\infty} \frac{x^k}{k!}\Big) = \sum_{k=0}^{\infty}\sum_{j=0}^{k} \frac{x^j}{j!}\frac{y^{k-j}}{(k-j)!}\\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j} x^j y^{k-j} = \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!}.\end{aligned}$$

(viii) We have $\exp(-x) = \frac{1}{\exp(x)}$ by part (vi), and so we compute

$$\lim_{x\to\infty} x^k \exp(-x) = \lim_{x\to\infty} \frac{x^k}{\sum_{j=0}^{\infty} \frac{x^j}{j!}} \le \lim_{x\to\infty} \frac{(k+1)!x^k}{x^{k+1}} = 0.$$

(ii) From parts (i) and (viii) we know that exp has an everywhere positive derivative. Thus, from Proposition 3.2.23 we know that exp is strictly monotonically increasing.

(iii) Clearly $\exp(x) > 0$ for all $x \in \mathbb{R}_{\ge 0}$. From part (vi) we have

$$\exp(x)\exp(-x) = \exp(0) = 1.$$

Therefore, for $x \in \mathbb{R}_{<0}$ we have $\exp(x) = \frac{1}{\exp(-x)} > 0$.

(iv) We have

$$\lim_{x\to\infty} \exp(x) = \lim_{x\to\infty} \sum_{j=0}^{\infty} \frac{x^j}{j!} \ge \lim_{x\to\infty} x = \infty.$$

(v) By parts (vi) and (iv) we have

$$\lim_{x\to-\infty} \exp(x) = \lim_{x\to\infty} \frac{1}{\exp(-x)} = 0.$$

(vii) Using part (vi) and the power series representation for exp we compute

$$\exp'(x) = \lim_{h\to 0} \frac{\exp(x+h) - \exp(x)}{h} = \lim_{h\to 0} \frac{\exp(x)(\exp(h) - 1)}{h} = \exp(x). \qquad \blacksquare$$

One of the reasons for the importance of the function exp in applications can be directly seen from property (vii). From this one can see that exp is the solution to the "initial value problem"

$$y'(x) = y(x), \quad y(0) = 1. \tag{3.18}$$

Most readers will recognise this as the differential equation governing a scalar process which exhibits "exponential growth." It turns out that many physical processes can be modelled, or approximately modelled, by such an equation, or by a suitable generalisation of such an equation. Indeed, one could use the solution of (3.18) as the *definition* of the function exp. However, to be rigorous, one would then be required to show that this equation has a unique solution; this is not altogether difficult, but does take one off topic a little. Such are the constraints imposed by rigour.

In Section 2.4.3 we defined the constant e by

$$\mathrm{e} = \sum_{j=0}^{\infty} \frac{1}{j!}.$$

From this we see immediately that $\mathrm{e} = \exp(1)$. To explore the relationship between the exponential function exp and the constant e, we first prove the following result, which recalls from Proposition 2.2.3 and the discussion immediately following it, the definition of x^q for $x \in \mathbb{R}_{>0}$ and $q \in \mathbb{Q}$.

3.6.3 Proposition ($\exp(x) = \mathrm{e}^x$) $\exp(x) = \sup\{\mathrm{e}^q \mid q \in \mathbb{Q},\ q < x\}$.

Proof First let us take the case where $x = q \in \mathbb{Q}$. Write $q = \frac{j}{k}$ for $j \in \mathbb{Z}$ and $k \in \mathbb{Z}_{>0}$. Then, by repeated application of part (vi) of Proposition 3.6.2 we have

$$\exp(q)^k = \exp(kq) = \exp(j) = \exp(j \cdot 1) = \exp(1)^j (\mathrm{e}^1)^j = \mathrm{e}^j.$$

By Proposition 2.2.3 this gives, by definition, $\exp(q) = \mathrm{e}^q$.

Now let $x \in \mathbb{R}$ and let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be a monotonically increasing sequence in $\mathbb{Q}$ such that $\lim_{j\to\infty} q_j = x$. By Theorem 3.1.3 we have $\exp(x) = \lim_{j\to\infty} \exp(q_j)$. By part (ii) of Proposition 3.6.2 the sequence $(\exp(q_j))_{j\in\mathbb{Z}_{>0}}$ is strictly monotonically increasing. Therefore, by Theorem 2.3.8,

$$\lim_{j\to\infty} \exp(q_j) = \lim_{j\to\infty} \mathrm{e}^{q_j} = \sup\{\mathrm{e}^q \mid q < x\},$$

as desired. ■

We shall from now on alternately use the notation e^x for $\exp(x)$, when this is more convenient.

3.6.2 The natural logarithmic function

From Proposition 3.6.2 we know that exp is a strictly monotonically increasing, continuous function. Therefore, by Theorem 3.1.30 we know that exp is an invertible function from $\mathbb{R}$ to image(exp). From parts (iii), (iv), and (v) of Proposition 3.6.2, as well as from Theorem 3.1.30 again, we know that image(exp) $= \mathbb{R}_{>0}$. This then leads to the following definition.

3.6.4 Definition (Natural logarithmic function) The ***natural logarithmic function***, denoted by $\log\colon \mathbb{R}_{>0} \to \mathbb{R}$, is the inverse of exp. •

We refer to Figure 3.16 for a depiction of the graph of log.

3.6.5 Notation (log versus ln) It is not uncommon to see the function that we denote by "log" written instead as "ln." In such cases, log is often used to refer to the base 10 logarithm (see Definition 3.6.13), since this convention actually sees much use in applications. However, we shall refer to the base 10 logarithm as $\log_{10}$. •

Now let us record the properties of log that follow immediately from its definition.

3.6.6 Proposition (Properties of the natural logarithmic function) *The natural logarithmic function enjoys the following properties:*

(i) log *is infinitely differentiable;*

(ii) log *is strictly monotonically increasing;*

(iii) $\log(x) = \int_1^x \frac{1}{\xi}\,d\xi$ *for all* $x \in \mathbb{R}_{>0}$*;*

(iv) $\lim_{x\to\infty} \log(x) = \infty$*;*

(v) $\lim_{x\downarrow 0} \log(x) = -\infty$*;*

(vi) $\log(xy) = \log(x) + \log(y)$ *for all* $x, y \in \mathbb{R}_{>0}$*;*

(vii) $\lim_{x\to\infty} x^{-k}\log(x) = 0$ *for all* $k \in \mathbb{Z}_{>0}$.

Proof (iii) From the Chain Rule and using the fact that $\log\circ\exp(x) = x$ for all $x \in \mathbb{R}$ we have

$$\log'(\exp(x)) = \frac{1}{\exp(x)} \quad\Longrightarrow\quad \log'(y) = \frac{1}{y}$$

for all $y \in \mathbb{R}_{>0}$. Using the fact that $\log(1) = 0$ (which follows since $\exp(0) = 1$), we then apply the Fundamental Theorem of Calculus, this being valid since $y \mapsto \frac{1}{y}$ is Riemann integrable on any compact interval in $\mathbb{R}_{>0}$, we obtain $\log(x) = \int_1^y \frac{1}{\eta}\,d\eta$, as desired.

(i) This follows from part (iii) using the fact that the function $x \mapsto \frac{1}{x}$ is infinitely differentiable on $\mathbb{R}_{>0}$.

(ii) This follows from Theorem 3.1.30.

(iv) We have

$$\lim_{x\to\infty} \log(x) = \lim_{y\to\infty} \log(\exp(y)) = \lim_{y\to\infty} y = \infty.$$

(v) We have

$$\lim_{x\downarrow 0} \log x = \lim_{y\to-\infty} \log(\exp(y)) = \lim_{y\to-\infty} y = -\infty.$$

(vi) For $x, y \in \mathbb{R}_{>0}$ write $x = \exp(a)$ and $y = \exp(b)$. Then

$$\log(xy) = \log(\exp(a)\exp(b)) = \log(\exp(a+b)) = a + b = \log(x) + \log(y).$$

(vii) We compute

$$\lim_{x\to\infty} \frac{\log x}{x^k} = \lim_{y\to\infty} \frac{\log\exp(y)}{\exp(y)^k} = \lim_{y\to\infty} \frac{y}{\exp(y)^k} \le \lim_{y\to\infty} \frac{y}{(1 + y + \frac{1}{2}y^2)^k} = 0. \qquad ■$$

3.6.3 Power functions and general logarithmic functions

For $x \in \mathbb{R}_{>0}$ and $q \in \mathbb{Q}$ we had defined, in and immediately following Proposition 2.2.3, x^q by $(x^{1/k})^j$ if $q = \frac{j}{k}$ for $j \in \mathbb{Z}$ and $k \in \mathbb{Z}_{>0}$. In this section we wish to extend this definition to x^y for $y \in \mathbb{R}$, and to explore the properties of the resulting function of both x and y.

3.6.7 Definition (Power function) If $a \in \mathbb{R}_{>0}$ then the function $\mathsf{P}_a \colon \mathbb{R} \to \mathbb{R}$ is defined by $\mathsf{P}_a(x) = \exp(x \log(a))$. If $a \in \mathbb{R}$ then the function $\mathsf{P}^a \colon \mathbb{R}_{>0} \to \mathbb{R}$ is defined by $\mathsf{P}^a(x) = \exp(a \log(x))$. •

Let us immediately connect this (when seen for the first time rather nonintuitive) definition to what we already know.

3.6.8 Proposition ($\mathbf{P_a(x) = a^x}$) $\mathrm{P_a(x) = \sup\{a^q \mid q \in \mathbb{Q},\ q < x\}}$.

Proof Let us first take $x = q \in \mathbb{Q}$ and write $q = \frac{j}{k}$ for $j \in \mathbb{Z}$ and $k \in \mathbb{Z}_{>0}$. We have

$$\exp(q \log(a))^k = \exp\Big(\frac{j}{k} \log(a)\Big)^k = \exp(j \log(a)) = \exp(\log(a))^j = a^j.$$

Therefore, by Proposition 2.2.3 we have

$$\exp(q \log(a)) = a^q.$$

Now let $x \in \mathbb{R}$ and let $(q_j)_{j \in \mathbb{Z}_{>0}}$ be a strictly monotonically increasing sequence in $\mathbb{Q}$ converging to x. Since exp and log are continuous, by Theorem 3.1.3 we have

$$\lim_{j \to \infty} \exp(q_j \log(a)) = \exp(x \log(a)).$$

As we shall see in Proposition 3.6.10, the function $x \mapsto \mathsf{P}_a(x)$ is strictly monotonically increasing. Therefore the sequence $(\exp(q_j \log(a)))_{j \in \mathbb{Z}_{>0}}$ is strictly monotonically increasing. Thus

$$\lim_{j \to \infty} \exp(q_j \log(a)) = \sup\{\mathsf{P}_a(q) \mid q \in \mathbb{Q},\ q < x\},$$

as desired. ■

Clearly we also have the following result.

3.6.9 Corollary ($\mathbf{P^a(x) = x^a}$) $\mathrm{P^a(x) = \sup\{x^q \mid q \in \mathbb{Q},\ q < a\}}$.

As with the exponential function, we will use the notation a^x for $\mathsf{P}_a(x)$ and x^a for $\mathsf{P}^a(x)$ when it is convenient to do so.

Let us now record some of the properties of the functions P_a and P^a that follow from their definition. When possible, we state the result using both the notation $\mathsf{P}_a(x)$ and a^x (or P^a and x^a).

3.6.10 Proposition (Properties of P_a) *For* $a \in \mathbb{R}_{>0}$*, the function* P_a *enjoys the following properties:*

(i) P_a *is infinitely differentiable;*

(ii) P_a *is strictly monotonically increasing when* $a > 1$*, is strictly monotonically decreasing when* $a < 1$*, and is constant when* $a = 1$*;*

(iii) $\mathsf{P}_a(x) = a^x > 0$ *for all* $x \in \mathbb{R}$*;*

(iv) $\lim_{x\to\infty} \mathsf{P}_a(x) = \lim_{x\to\infty} a^x = \begin{cases} \infty, & a > 1, \\ 0, & a < 1, \\ 1, & a = 1; \end{cases}$

(v) $\lim_{x\to-\infty} \mathsf{P}_a(x) == \lim_{x\to-\infty} a^x = \begin{cases} 0, & a > 1, \\ \infty, & a < 1, \\ 1, & a = 1; \end{cases}$

(vi) $\mathsf{P}_a(x+y) = a^{x+y} = a^x a^y = \mathsf{P}_a(x)\mathsf{P}_a(y)$*;*

(vii) $\mathsf{P}_a'(x) = \log(a)\mathsf{P}_a(x)$*;*

(viii) *if* $a > 1$ *then* $\lim_{x\to\infty} x^k \mathsf{P}_a(-x) = \lim_{x\to\infty} x^k a^{-x} = 0$ *for all* $k \in \mathbb{Z}_{>0}$*;*

(ix) *if* $a < 1$ *then* $\lim_{x\to\infty} x^k \mathsf{P}_a(x) = \lim_{x\to\infty} x^k a^x = 0$ *for all* $k \in \mathbb{Z}_{>0}$*.*

Proof (i) Define $f, g\colon \mathbb{R} \to \mathbb{R}$ and $f(x) = x\log(a)$ and $g(x) = \exp(x)$. Then $\mathsf{P}_a = g \circ f$, and so is the composition of infinitely differentiable functions. This part of the result follows from Theorem 3.2.13.

(ii) Let $x_1 < x_2$. If $a > 1$ then $\log(a) > 0$ and so

$$x_1 \log(a) < x_2 \log(a) \quad\Longrightarrow\quad \exp(x_1 \log(a)) < \exp(x_2 \log(a))$$

since exp is strictly monotonically increasing. If $a < 1$ then $\log(a) < 0$ and so

$$x_1 \log(a) > x_2 \log(a) \quad\Longrightarrow\quad \exp(x_1 \log(a)) > \exp(x_2 \log(a)),$$

again since exp is strictly monotonically increasing. For $a = 1$ we have $\log(a) = 0$ so $\mathsf{P}_a(x) = 1$ for all $x \in \mathbb{R}$.

(iii) This follows since $\operatorname{image}(\exp) \subseteq \mathbb{R}_{>0}$.

(iv) For $a > 1$ we have

$$\lim_{x\to\infty} \mathsf{P}_a(x) = \lim_{x\to\infty} \exp(x\log(a)) = \lim_{y\to\infty} \exp(y) = \infty,$$

and for $a < 1$ we have

$$\lim_{x\to\infty} \mathsf{P}_a(x) = \lim_{x\to\infty} \exp(x\log(a)) = \lim_{y\to-\infty} \exp(y) = 0.$$

For $a = 1$ the result is clear since $\mathsf{P}_1(x) = 1$ for all $x \in \mathbb{R}$.

(v) For $a > 1$ we have

$$\lim_{x\to-\infty} \mathsf{P}_a(x) = \lim_{x\to-\infty} \exp(x\log(a)) = \lim_{y\to-\infty} \exp(y) = 0,$$

and for $a < 1$ we have

$$\lim_{x\to-\infty} \mathsf{P}_a(x) = \lim_{x\to-\infty} \exp(x\log(a)) = \lim_{y\to\infty} \exp(y) = \infty.$$

Again, for $a = 1$ the result is obvious.

(vi) We have

$$\mathsf{P}_a(x+y) = \exp((x+y)\log(a)) = \exp(x\log(a))\exp(y\log(a)) = \mathsf{P}_a(x)\mathsf{P}_a(y).$$

(vii) With f and g as in part (i), and using Theorem 3.2.13, we compute

$$\mathsf{P}_a'(x) = g'(f(x))f'(x) = \exp(x\log(a))\log(a) = \log(a)\mathsf{P}_a(x).$$

(viii) We compute

$$\lim_{x\to\infty} x^k \mathsf{P}_a(-x) = \lim_{x\to\infty} x^k \exp(-x\log(a)) = \lim_{y\to\infty}\Big(\frac{y}{\log(a)}\Big)^k \exp(-y) = 0,$$

using part (viii) of Proposition 3.6.2.

(ix) We have

$$\lim_{x\to\infty} x^k \mathsf{P}_a(x) = \lim_{x\to\infty} x^k \exp((-x)(-\log(a))) = 0$$

since $\log(a) < 0$. ■

3.6.11 Proposition (Properties of P^a) *For $a \in \mathbb{R}$, the function P^a enjoys the following properties:*

(i) P^a *is infinitely differentiable;*

(ii) P^a *is strictly monotonically increasing;*

(iii) $\mathsf{P}^a(x) = x^a > 0$ *for all* $x \in \mathbb{R}_{>0}$*;*

(iv) $\lim_{x\to\infty} \mathsf{P}^a(x) = \lim_{x\to\infty} x^a = \begin{cases} \infty, & a > 0, \\ 0, & a < 0, \\ 1, & a = 0; \end{cases}$

(v) $\lim_{x\downarrow 0} \mathsf{P}^a(x) = \lim_{x\downarrow 0} x^a = \begin{cases} 0, & a > 0, \\ \infty, & a < 0, \\ 1, & a = 0; \end{cases}$

(vi) $\mathsf{P}^a(xy) = (xy)^a = x^a y^a = \mathsf{P}^a(x)\mathsf{P}^a(y)$*;*

(vii) $(\mathsf{P}^a)'(x) = a\mathsf{P}^{a-1}(x)$*.*

Proof (i) Define $f\colon \mathbb{R}_{>0} \to \mathbb{R}$, $g\colon \mathbb{R} \to \mathbb{R}$, and $h\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = \log(x)$, $g(x) = ax$, and $h(x) = \exp(x)$. Then $\mathsf{P}^a = h \circ g \circ f$. Since each of f, g, and h is infinitely differentiable, then so too is P^a by Theorem 3.2.13.

(ii) Let $x_1, x_2 \in \mathbb{R}_{>0}$ satisfy $x_1 < x_2$. Then

$$\mathsf{P}^a(x_1) = \exp(a\log(x_1)) < \exp(a\log(x_2)) = \mathsf{P}^a(x_2)$$

using the fact that both log and exp are strictly monotonically increasing.

(iii) This follows since image(exp) $\subseteq \mathbb{R}_{>0}$.

(iv) For $a > 0$ we have

$$\lim_{x\to\infty} \mathsf{P}^a(x) = \lim_{x\to\infty} \exp(a\log(x)) = \lim_{y\to\infty} \exp(y) = \infty,$$

and for $a < 0$ we have

$$\lim_{x\to\infty} \mathsf{P}^a(x) = \lim_{x\to\infty} \exp(a\log(x)) = \lim_{y\to-\infty} \exp(y) = 0.$$

For $a = 0$ we have $\mathsf{P}^a(x) = 1$ for all $x \in \mathbb{R}_{>0}$.

(v) For $a > 0$ we have

$$\lim_{x\downarrow 0} \mathsf{P}^a(x) = \lim_{x\downarrow 0} \exp(a\log(x)) = \lim_{y\to-\infty} \exp(y) = 0,$$

and for $a < 0$ we have

$$\lim_{x\downarrow 0} \mathsf{P}^a(x) = \lim_{x\downarrow 0} \exp(a\log(x)) = \lim_{y\to\infty} \exp(y) = \infty.$$

For $a = 1$, the result is trivial again.

(vi) We have

$$\mathsf{P}^a(xy) = \exp(a\log(xy)) = \exp(a(\log(x)+\log(y))) = \exp(a\log(x))\exp(a\log(y)) = \mathsf{P}^a(x)\mathsf{P}^a(y).$$

(vii) With f, g, and h as in part (i), and using the Chain Rule, we have

$$\begin{aligned}(\mathsf{P}^a)'(x) &= h'(g(f(x)))g'(f(x))f'(x) = a\exp(a\log(x))\tfrac{1}{x}\\ &= a\exp(a\log(x))\exp(-1\log(x)) = a\exp((a-1)\log(x)) = a\mathsf{P}^{a-1}(x),\end{aligned}$$

as desired, using part (vi) of Proposition 3.6.10. ■

The following result is also sometimes useful.

3.6.12 Proposition (Property of $\mathsf{P}_x(x^{-1})$) $\lim_{x\to\infty} \mathsf{P}_x(x^{-1}) = \lim_{x\to\infty} x^{1/x} = 1$.

Proof We have

$$\lim_{x\to\infty} \mathsf{P}_x(x^{-1}) = \lim_{x\to\infty} \exp(x^{-1}\log(x)) = \lim_{y\to 0} \exp(y) = 1,$$

using part (vii) of Proposition 3.6.6. ■

Now we turn to the process of inverting the power function. For the exponential function we required that $\log(e^x) = x$. Thus, if our inverse of P_a is denoted (for the moment) by f_a, then we expect that $f_a(a^x) = x$. This definition clearly has difficulties when $a = 1$, reflecting the fact that P_1 is not invertible. In all other case, since P_a is continuous, and either strictly monotonically increasing or strictly monotonically decreasing, we have the following definition, using Theorem 3.1.30.

3.6.13 Definition (Arbitrary base logarithm) For $a \in \mathbb{R}_{>0}\setminus\{1\}$, the function $\log_a\colon \mathbb{R}_{>0} \to \mathbb{R}$, called the ***base a logarithmic function***, is the inverse of P_a. When $a = 10$ we simply write $\log_{10} = \log$. •

The following result relates the logarithmic function for an arbitrary base to the natural logarithmic function.

3.6.14 Proposition (Characterisation of $\log_{\mathrm{a}}$) $\log_{\mathrm{a}}(\mathrm{x}) = \dfrac{\log(\mathrm{x})}{\log(\mathrm{a})}$.

Proof Let $x \in \mathbb{R}_{>0}$ and write $x = a^y$ for some $y \in \mathbb{R}$. First suppose that $y \neq 0$. Then we have $\log(x) = y\log(a)$ and $\log_a(x) = y$, and the result follows by eliminating y from these two expressions. When $y = 0$ we have $x = a = a^1$. Therefore, $\log_a(x) = 1 = \frac{\log(x)}{\log(a)}$. ■

With this result we immediately have the following generalisation of Proposition 3.6.6. We leave the trivial checking of the details to the reader.

3.6.15 Proposition (Properties of $\log_{\mathrm{a}}$) *For* $\mathrm{a} \in \mathbb{R}_{>0} \setminus \{1\}$, *the function* $\log_{\mathrm{a}}$ *enjoys the following properties:*

(i) $\log_{\mathrm{a}}$ *is infinitely differentiable;*

(ii) $\log_{\mathrm{a}}$ *is strictly monotonically increasing when* $\mathrm{a} > 1$ *and is strictly monotonically decreasing when* $\mathrm{a} < 1$*;*

(iii) $\log_{\mathrm{a}}(\mathrm{x}) = \frac{1}{\log(\mathrm{a})}\int_1^{\mathrm{x}} \frac{1}{\xi}\,\mathrm{d}\xi$ *for all* $\mathrm{x} \in \mathbb{R}_{>0}$*;*

(iv) $\lim_{\mathrm{x}\to\infty} \log_{\mathrm{a}}(\mathrm{x}) = \begin{cases} \infty, & \mathrm{a} > 1, \\ -\infty, & \mathrm{a} < 1; \end{cases}$

(v) $\lim_{\mathrm{x}\downarrow 0} \log_{\mathrm{a}}(\mathrm{x}) = \begin{cases} -\infty, & \mathrm{a} > 1, \\ \infty, & \mathrm{a} < 1; \end{cases}$

(vi) $\log_{\mathrm{a}}(\mathrm{xy}) = \log_{\mathrm{a}}(\mathrm{x}) + \log_{\mathrm{a}}(\mathrm{y})$ *for all* $\mathrm{x}, \mathrm{y} \in \mathbb{R}_{>0}$*;*

(vii) $\lim_{\mathrm{x}\to\infty} \mathrm{x}^{-\mathrm{k}} \log_{\mathrm{a}}(\mathrm{x}) = 0$ *for all* $\mathrm{k} \in \mathbb{Z}_{>0}$.

3.6.4 Trigonometric functions

Next we turn to describing the standard trigonometric functions. These functions are perhaps most intuitively introduced in terms of the concept of "angle" in plane geometry. However, to really do this properly would, at this juncture, require a significant expenditure of effort. Therefore, we define the trigonometric functions by their power series expansion, and then proceed to show that they have the expected properties. In the course of our treatment we will also see that the constant π introduced in Section 2.4.3 has the anticipated relationships to the trigonometric functions. Convenience in this section forces us to make a fairly serious logical jump in the presentation. While all constructions and theorems are stated in terms of real numbers, in the proofs we use complex numbers rather heavily.

3.6.16 Definition (sin and cos) The ***sine function***, denoted by $\sin\colon \mathbb{R} \to \mathbb{R}$, and the ***cosine function***, denoted by $\cos\colon \mathbb{R} \to \mathbb{R}$, are defined by

$$\sin(x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1} x^{2j-1}}{(2j-1)!}, \quad \cos(x) = \sum_{j=0}^{\infty} \frac{(-1)^j x^{2j}}{(2j)!},$$

respectively. •

In Figure 3.17 we show the graphs of the functions sin and cos.

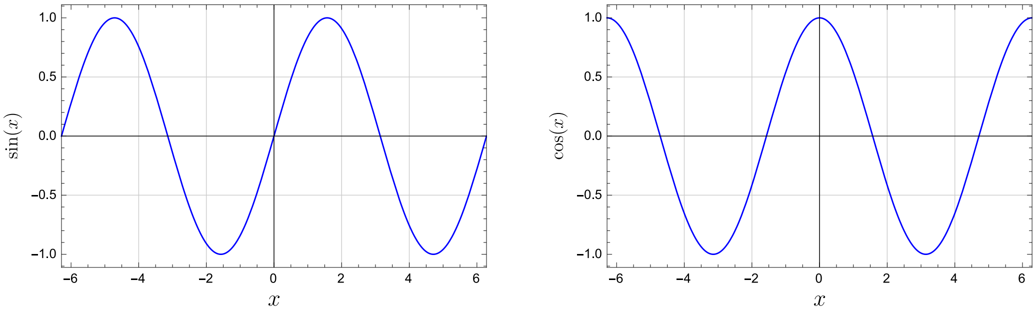


Figure 3.17 The functions sin (left) and cos (right)

3.6.17 Notation Following normal conventions, we shall frequently write $\sin x$ and $\cos x$ rather than the more correct $\sin(x)$ and $\cos(x)$. •

An application of Proposition 2.4.15 and Theorem **??** shows that the power series expansions for sin and cos are, in fact, convergent for all x, and so the functions are indeed defined with domain $\mathbb{R}$.

First we prove the existence of a number having the property that we know π to possess. In fact, we construct the number $\frac{\pi}{2}$, where π is as given in Section 2.4.3.

3.6.18 Theorem (Construction of π) *There exists a positive real number* p_0 *such that*

$$\mathrm{p}_0 = \inf\{\mathrm{x} \in \mathbb{R}_{>0} \mid \cos(\mathrm{x}) = 0\}.$$

Moreover, $\mathrm{p}_0 = \frac{\pi}{2}$.

Proof First we record the derivative properties for sin and cos.

1 Lemma *The functions* sin *and* cos *are infinitely differentiable and satisfy* $\sin' = \cos$ *and* $\cos' = -\sin$.

Proof This follows directly from Proposition **??** where it is shown that convergent power series can be differentiated term-by-term. ▼

Let us now perform some computations using complex variables that will be essential to many of the proofs in this section. We suppose the reader to be acquainted with the necessary elementary facts about complex numbers. The next observation is the most essential along these lines. We denote $\mathbb{S}^{\mathbb{C}}_1 = \{z \in \mathbb{C} \mid |z| = 1\}$, and recall that all points in $z \in \mathbb{S}^1_{\mathbb{C}}$ can be written as $z = \mathrm{e}^{\mathrm{i}x}$ for some $x \in \mathbb{R}$, and that, conversely, for any $x \in \mathbb{R}$ we have $\mathrm{e}^{\mathrm{i}x} \in \mathbb{S}^1_{\mathbb{C}}$.

2 Lemma $\mathrm{e}^{\mathrm{ix}} = \cos(\mathrm{x}) + \mathrm{i}\sin(\mathrm{x})$.

Proof This follows immediately from the $\mathbb{C}$-power series for the complex exponential function:

$$\mathrm{e}^z = \sum_{j=0}^{\infty} \frac{x^j}{j!}.$$

Substituting $z = \mathrm{i}x$, using the fact that $\mathrm{i}^{2j} = (-1)^j$ for all $j \in \mathbb{Z}_{>0}$, and using Proposition 2.4.30, we get the desired result. ▼

From the preceding lemma we then know that $\cos(x) = \mathrm{Re}(\mathrm{e}^{\mathrm{i}x})$ and that $\sin(x) = \mathrm{Im}(\mathrm{e}^{\mathrm{i}x})$. Therefore, since $\mathrm{e}^{\mathrm{i}x} \in \mathbb{S}^1_{\mathbb{C}}$, we have

$$\cos(x)^2 + \sin(x)^2 = 1. \tag{3.19}$$

Let us show that the set $\{x \in \mathbb{R}_{>0} \mid \cos(x) = 0\}$ is nonempty. Suppose that it is empty. Since $\cos(0) = 1$ and since cos is continuous, it must therefore be the case (by the Intermediate Value Theorem) that $\cos(x) > 0$ for all $x \in \mathbb{R}$. Therefore, by Lemma 1, $\sin'(x) > 0$ for all $x \in \mathbb{R}$, and so sin is strictly monotonically increasing by Proposition 3.2.23. Therefore, since $\sin(0) = 0$, $\sin(x) > 0$ for $x > 0$. Therefore, for $x_1, x_2 \in \mathbb{R}_{>0}$ satisfying $x_1 < x_2$, we have

$$\sin(x_1)(x_2 - x_1) < \int_{x_1}^{x_2} \sin(x)\,\mathrm{d}x = \cos(x_2) - \cos(x_1) \leq 2,$$

where we have used the fact that sin is strictly monotonically increasing, Lemma 1, the Fundamental Theorem of Calculus, and (3.19). We thus have arrive at the contradiction that $\limsup_{x_2 \to \infty} \sin(x_1)(x_2 - x_1) \leq 2$.

Since cos is continuous, the set $\{x \in \mathbb{R}_{>0} \mid \cos(x) = 0\}$ is closed. Therefore, $\inf\{x \in \mathbb{R}_{>0} \mid \cos(x) = 0\}$ is contained in this set, and this gives the existence of p_0. Note that, by (3.19), $\sin(p_0) \in \{-1, 1\}$. Since $\sin(0) = 0$ and since $\sin(x) = \cos(x) > 0$ for $x \in [0, p_0)$, we must have $\sin(p_0) = 1$.

The following property of p_0 will also be important.

3 Lemma $\cos(\frac{p_0}{2}) = \sin(\frac{p_0}{2}) = \frac{1}{\sqrt{2}}$.

Proof Let $x_0 = \cos(\frac{p_0}{2})$, $y_0 = \sin(\frac{p_0}{2})$, and $z_0 = x_0 + \mathrm{i}y_0$. Then, using Proposition **??**,

$$(\mathrm{e}^{\mathrm{i}\frac{p_0}{2}})^2 = \mathrm{e}^{\mathrm{i}p_0} = \mathrm{i}$$

since $\cos(p_0) = 0$ and $\sin(p_0) = 1$. Thus

$$(\mathrm{e}^{\mathrm{i}\frac{p_0}{2}})^4 = \mathrm{i}^2 = -1,$$

again using Proposition **??**. Using the definition of complex multiplication we also have

$$(\mathrm{e}^{\mathrm{i}\frac{p_0}{2}})^4 = (x_0 + \mathrm{i}y_0)^4 = x_0^4 - 6x_0^2y_0^2 + y_0^4 + 4\mathrm{i}x_0y_0(x_0^2 - y_0^2).$$

Thus, in particular, $x_0^2 - y_0^2 = 0$. Combining this with $x_0^2 + y_0^2 = 1$ we get $x_0^2 = y_0^2 = \frac{1}{2}$. Since both x_0 and y_0 are positive by virtue of $\frac{p_0}{2}$ lying in $(0, p_0)$, we must have $x_0 = y_0 = \frac{1}{\sqrt{2}}$, as claimed. ▼

Now we show, through a sequence of seemingly irrelevant computations, that $p_0 = \frac{\pi}{2}$. Define the function $\tan\colon (-p_0, p_0) \to \mathbb{R}$ by $\tan(x) = \frac{\sin(x)}{\cos(x)}$, noting that tan is well-defined since $\cos(-x) = \cos(x)$ and since $\cos(x) > 0$ for $x \in [0, p_0)$. We claim that tan is continuous and strictly monotonically increasing. We have, using the quotient rule,

$$\tan'(x) = \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2}.$$

Thus $\tan'(x) > 0$ for all $x \in (-p_0, p_0)$, and so tan is strictly monotonically increasing by Proposition 3.2.23. Since $\sin(p_0) = 1$ and (since $\sin(-x) = -\sin(x)$) since $\sin(-p_0) = -1$, we have

$$\lim_{x \uparrow p_0} \tan(x) = \infty, \quad \lim_{x \downarrow p_0} \tan(x) = -\infty.$$

This shows that tan is an invertible and differentiable mapping from $(-p_0, p_0)$ to $\mathbb{R}$. Moreover, since $\tan'$ is nowhere zero, the inverse, denoted by $\tan^{-1} \colon \mathbb{R} \to (-p_0, p_0)$, is also differentiable and the derivative of its inverse is given by

$$(\tan^{-1})'(x) = \frac{1}{\tan'(\tan^{-1}(x))},$$

as per Theorem 3.2.24. We further claim that

$$(\tan^{-1})'(x) = \frac{1}{1+x^2}.$$

Indeed, our above arguments show that $(\tan^{-1})'(x) = (\cos(\tan^{-1}(x)))^2$. If $y = \tan^{-1}(x)$ then

$$\frac{\sin(y)}{\cos(y)} = x.$$

Since $\sin(y) > 0$ for $y \in (0, p_0)$, we have $\sin(y) = \sqrt{1 - \cos(y)}$ by (3.19). Therefore,

$$\frac{1 - \cos(y)^2}{\cos(y)^2} = x^2 \quad \Longrightarrow \quad \cos(y)^2 = \frac{1}{1+x^2}$$

as desired.

By the Fundamental Theorem of Calculus we then have

$$\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = \tan^{-1}(1) - \tan^{-1}(0).$$

Since $\tan^{-1}(1) = \frac{p_0}{2}$ by Lemma 3 above and since $\tan^{-1}(0) = 0$ (and using part (v) of Proposition 3.6.19 below), we have

$$\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = \frac{p_0}{2}. \tag{3.20}$$

Now recall from Example **??**–**??** that we have

$$\frac{1}{1+x^2} = \sum_{j=0}^{\infty} (-1)^j x^{2j},$$

with the series converging uniformly on any compact subinterval of $(-1, 1)$. Therefore, by Proposition **??**, for $\epsilon \in (0, 1)$ we have

$$\begin{aligned}
\int_0^{1-\epsilon} \frac{1}{1+x^2}\,\mathrm{d}x &= \int_0^{1-\epsilon} \sum_{j=0}^{\infty} (-1)^j x^{2j}\,\mathrm{d}x \\
&= \sum_{j=0}^{\infty} (-1)^j \int_0^{1-\epsilon} x^{2j}\,\mathrm{d}x \\
&= \sum_{j=0}^{\infty} (-1)^j \frac{(1-\epsilon)^{2j+1}}{2j+1}.
\end{aligned}$$

The following technical lemma will allow us to conclude the proof.

4 Lemma $\lim_{\epsilon\downarrow 0}\sum_{\mathrm{j}=0}^{\infty}(-1)^{\mathrm{j}}\frac{(1-\epsilon)^{2\mathrm{j}+1}}{2\mathrm{j}+1}=\sum_{\mathrm{j}=0}\frac{(-1)^{\mathrm{j}}}{2\mathrm{j}+1}.$

Proof By the Alternating Test, the series $\sum_{j=0}^{\infty}(-1)^j\frac{(1-\epsilon)^{2j+1}}{2j+1}$ converges for $\epsilon\in[0,2]$. Define $f\colon[0,2]\to\mathbb{R}$ by

$$f(x)=\sum_{j=0}^{\infty}(-1)^{j+1}\frac{(x-1)^{2j+1}}{2j+1}$$

and define $g\colon[-1,1]\to\mathbb{R}$ by

$$g(x)=\sum_{j=0}^{\infty}(-1)^{j+1}\frac{x^{2j+1}}{2j+1}$$

so that $f(x)=g(x-1)$. Since g is defined by a $\mathbb{R}$-convergent power series, by Corollary **??** g is continuous. In particular,

$$g(-1)=\lim_{x\downarrow -1}\sum_{j=0}^{\infty}(-1)^{j+1}\frac{x^{2j+1}}{2j+1}.$$

From this it follows that

$$f(0)=\lim_{x\downarrow 0}\sum_{j=0}^{\infty}(-1)^{j+1}\frac{(x-1)^{2j+1}}{2j+1},$$

which is the result. ▼

Combining this with (3.20) we have

$$\frac{p_0}{2}=\lim_{\epsilon\downarrow 0}\int_0^{1-\epsilon}\frac{1}{1+x^2}\,\mathrm{d}x=\lim_{\epsilon\downarrow 0}\sum_{j=0}^{\infty}(-1)^j\frac{(1-\epsilon)^{2j+1}}{2j+1}=\sum_{j=0}\frac{(-1)^j}{2j+1}=\frac{\pi}{4},$$

using the definition of π in Definition 2.4.20. ■

Now that we have on hand a reasonable characterisation of π, we can proceed to state the familiar properties of sin and cos.

3.6.19 Proposition (Properties of sin and cos) *The functions* sin *and* cos *enjoy the following properties:*

(i) sin *and* cos *are infinitely differentiable, and furthermore satisfy* $\sin'=\cos$ *and* $\cos'=-\sin$*;*

(ii) $\sin(-\mathrm{x})=\sin(\mathrm{x})$ *and* $\cos(-\mathrm{x})=\cos(\mathrm{x})$ *for all* $\mathrm{x}\in\mathbb{R}$*;*

(iii) $\sin(\mathrm{x})^2+\cos(\mathrm{x})^2=1$ *for all* $\mathrm{x}\in\mathbb{R}$*;*

(iv) $\sin(\mathrm{x}+2\pi)=\sin(\mathrm{x})$ *and* $\cos(\mathrm{x}+2\pi)=\cos(\mathrm{x})$ *for all* $\mathrm{x}\in\mathbb{R}$*;*

(v) the map

$$[0,2\pi)\ni\mathrm{x}\mapsto(\cos(\mathrm{x}),\sin(\mathrm{x}))\in\{(\mathrm{x},\mathrm{y})\in\mathbb{R}^2\mid\ \mathrm{x}^2+\mathrm{y}^2=1\}$$

is a bijection.

Proof (i) This was proved as Lemma 1 in the proof of Theorem 3.6.18.

(ii) This follows immediately from the $\mathbb{R}$-power series for sin and cos.

(iii) This was proved as (3.19) in the course of the proof of Theorem 3.6.18.

(iv) Since $e^{i\frac{\pi}{2}} = i$ by Theorem 3.6.18, we use Proposition **??** to deduce

$$e^{2\pi i} = (e^{i\frac{\pi}{2}})^4 = i^4 = 1.$$

Again using Proposition **??** we then have

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z$$

for all $z \in \mathbb{C}$. Therefore, for $x \in \mathbb{R}$, we have

$$\cos(x + 2\pi) + i\sin(x + 2\pi) = e^{i(x+2\pi)} = e^{ix} = \cos(x) + i\sin(x),$$

which gives the result.

(v) Denote $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, and note that, if we make the standard identification of $\mathbb{C}$ with $\mathbb{R}^2$ (as we do), then $\mathbb{S}^1_{\mathbb{C}}$ (see the proof of Theorem 3.6.18) becomes identified with $\mathbb{S}^1$, with the identification explicitly being $x + iy \mapsto (x, y)$. Thus the result we are proving is equivalent to the assertion that the map

$$f\colon [0, 2\pi) \ni x \mapsto e^{ix} \in \mathbb{S}^1_{\mathbb{C}}$$

is a bijection. This is what we will prove. By part (iii), this map is well-defined in the sense that it actually does take values in $\mathbb{S}^1_{\mathbb{C}}$. Suppose that $e^{ix_1} = e^{ix_2}$ for distinct points $x_1, x_2 \in [0, 2\pi)$, and suppose for concreteness that $x_1 < x_2$. Then $x_2 - x_1 \in (0, 2\pi)$, and $\frac{1}{4}(x_2 - x_1) \in (0, \frac{\pi}{2})$. We then have

$$e^{ix_1} = e^{ix_2} \quad\Longrightarrow\quad e^{i(x_2-x_1)} = 1 \quad\Longrightarrow\quad (e^{i\frac{1}{4}(x_2-x_1)})^4 = 1.$$

Let $e^{i\frac{1}{4}(x_2-x_1)} = \xi + i\eta$. Since $\frac{1}{4}(x_2 - x_1) \in (0, \frac{\pi}{2})$, we saw during the course of the proof of Theorem 3.6.18 that $\xi, \eta \in (0, 1)$. We then use the definition of complex multiplication to compute

$$(e^{i\frac{1}{4}(x_2-x_1)})^4 = \xi^4 - 6\xi^2\eta^2 + \eta^4 + 4i\xi\eta(\xi^2 - \eta^2).$$

Since $(e^{i\frac{1}{4}(x_2-x_1)})^4 = 1$ is real, we conclude that $\xi^2 - \eta^2 = 0$. Combining this with $\xi^2 + \eta^2 = 1$ gives $\xi^2 = \eta^2 = \frac{1}{2}$. Since both ξ and η are positive we have $\xi = \eta = \frac{1}{\sqrt{2}}$. Substituting this into the above expression for $(e^{i\frac{1}{4}(x_2-x_1)})^4$ gives $(e^{i\frac{1}{4}(x_2-x_1)})^4 = -1$. Thus we arrive at a contradiction, and it cannot be the case that $e^{ix_1} = e^{ix_2}$ for distinct $x_1, x_2 \in [0, 2\pi)$. Thus f is injective.

To show that f is surjective, we let $z = x + iy \in \mathbb{S}^1_{\mathbb{C}}$, and consider four cases.

1. $x, y \geq 0$: Since cos is monotonically decreasing from 1 to 0 on $[0, \frac{\pi}{2}]$, there exists $\theta \in [0, \frac{\pi}{2}]$ such that $\cos(\theta) = x$. Since $\sin(\theta)^2 = 1 - \cos(\theta)^2 = 1 - x^2 = y^2$, and since $\sin(\theta) \geq 0$ for $\theta \in [0, \frac{\pi}{2}]$, we conclude that $\sin(\theta) = y$. Thus $z = e^{i\theta}$.
2. $x \geq 0$ and $y \leq 0$: Let $\xi = x$ and $\eta = -y$ so that $\xi, \eta \geq 0$. From the preceding case we deduce the existence of $\phi \in [0, \frac{\pi}{2}]$ such that $e^{i\phi} = \xi + i\eta$. Thus $\cos(\phi) = x$ and $\sin(\phi) = -y$. By part (ii) we then have $\cos(-\phi) = x$ and $\sin(-\phi) = y$, and we note that $-\phi \in [-\frac{\pi}{2}, 0]$. Define

$$\theta = \begin{cases} 2\pi - \phi, & \phi \in (0, \frac{\pi}{2}], \\ 0, & \phi = 0. \end{cases}$$

By part (iv) we then have $\cos(\theta) = x$ and $\sin(\theta) = y$, and that $\theta \in [\frac{3\pi}{2}, 2\pi)$ if $\phi \in (0, \frac{\pi}{2}]$.

3. $x \leq 0$ and $y \geq 0$: Let $\xi = -x$ and $\eta = y$ si that $\xi, \eta \geq 0$. As in the first case we have $\phi \in [0, \frac{\pi}{2}]$ such that $\cos(\phi) = \xi$ and $\sin(\phi) = \eta$. We then have $-\cos(\phi) = x$ and $\sin(\phi) = y$. Next define $\theta = \pi - \phi$ and note that

$$e^{i\theta} = e^{i\pi}e^{-i\phi} = -(\cos(\phi) - i\sin(\phi)) = -\cos(\phi) + i\sin(\phi) = x + iy,$$

as desired.

4. $x \leq 0$ and $y \leq 0$: Take $\xi = -x$ and $\eta = -y$ so that $\xi, \eta \geq 0$. As in the first case, we have $\phi \in [0, \frac{\pi}{2}]$ such that $\cos(\phi) = \xi = -x$ and $\sin(\phi) = \eta = -y$. Then, taking $\theta = \pi + \phi$, we have

$$e^{i\theta} = e^{i\pi}e^{i\phi} = -(\cos(\phi) + i\sin(\phi)) = x + iy,$$

as desired. ■

From the basic construction of sin and cos that we give, and the properties that follow directly from this construction, there is of course a great deal that one can proceed to do; the resulting subject is broadly called “trigonometry.” Rigorous proofs of many of the facts of basic trigonometry follow easily from our constructions here, particularly since we give the necessary properties, along with a rigorous definition, of π. We do assume that the reader has an acquaintance with trigonometry, as we shall use certain of these facts without much ado.

The reciprocals of sin and cos are sometimes used. Thus we define $\csc\colon (0, 2\pi) \to \mathbb{R}$ and $\sec\colon (-\pi, \pi) \to \mathbb{R}$ by $\csc(x) = \frac{1}{\sin(x)}$ and $\sec(x) = \frac{1}{\cos(x)}$. These are the ***cosecant*** and ***secant*** functions, respectively. One can verify that the restrictions of csc and sec to $(0, \frac{\pi}{2})$ are bijective. In Figure 3.18

One useful and not perfectly standard construction is the following. Define $\tan\colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ by $\tan(x) = \frac{\sin(x)}{\cos(x)}$, noting that the definition makes sense since $\cos(x) > 0$ for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. In Figure 3.19 we depict the graph of tan and its inverse $\tan^{-1}$. During the course of the proof of Theorem 3.6.18 we showed that the function tan had the following properties.

3.6.20 Proposition (Properties of tan) *The function* tan *enjoys the following properties:*

(i) tan *is infinitely differentiable;*

(ii) tan *is strictly monotonically increasing;*

(iii) the inverse of tan, *denoted by* $\tan^{-1}\colon \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ *is infinitely differentiable.*

It turns out to be useful to extend the definition of $\tan^{-1}$ to $(-\pi, \pi]$ by defining the function $\operatorname{atan}\colon \mathbb{R}^2 \setminus \{(0,0)\} \to (-\pi, \pi]$ by

$$\operatorname{atan}(x, y) = \begin{cases} \tan^{-1}(\frac{y}{x}), & x > 0, \\ \pi - \tan^{-1}(\frac{y}{x}), & x < 0, \\ \frac{\pi}{2}, & x = 0,\ y > 0, \\ -\frac{\pi}{2}, & x = 0,\ y < 0. \end{cases}$$

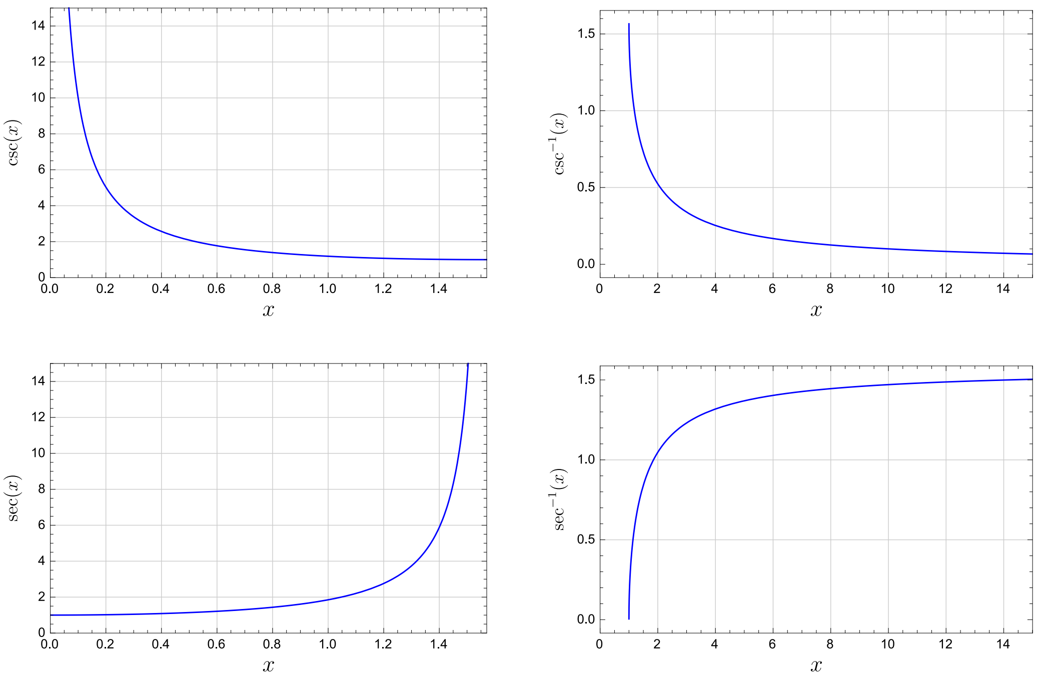


Figure 3.18 Cosecant and its inverse (top) and secant and its inverse (bottom) on $(0, \frac{\pi}{2})$

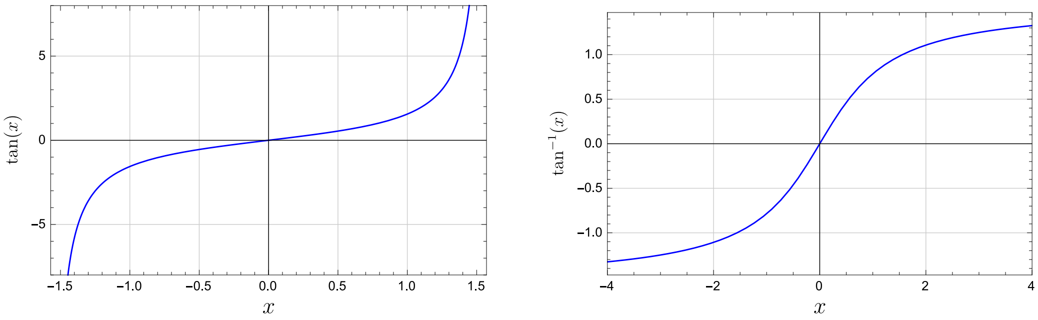


Figure 3.19 The function tan (left) and its inverse $\tan^{-1}$ (right)

As we shall see in ***missing stuff*** when we discuss the geometry of the complex plane, this function returns that angle of a point (x, y) measured from the positive x-axis.

3.6.5 Hyperbolic trigonometric functions

In this section we shall quickly introduce the hyperbolic trigonometric functions. Just why these functions are called "trigonometric" is only best seen in the setting of $\mathbb{C}$-valued functions in ***missing stuff***.

3.6.21 Definition (sinh and cosh) The ***hyperbolic sine function***, denoted by $\sinh\colon \mathbb{R} \to \mathbb{R}$, and the ***hyperbolic cosine function***m denoted by $\cosh\colon \mathbb{R} \to \mathbb{R}$, are defined by

$$\sinh(x) = \sum_{j=1}^{\infty} \frac{x^{2j-1}}{(2j-1)!}, \quad \cosh(x) = \sum_{j=0}^{\infty} \frac{x^{2j}}{(2j)!},$$

respectively. •

In Figure 3.20 we depict the graphs of sinh and cosh.

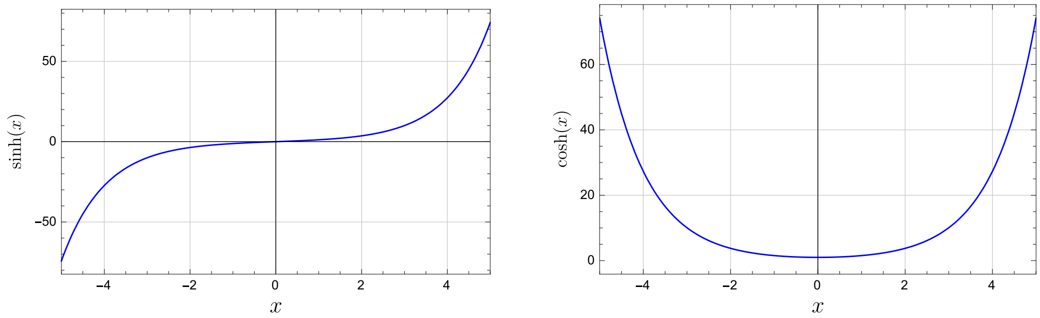


Figure 3.20 The functions sinh (left) and cosh (right)

As with sin and cos, an application of Proposition 2.4.15 and Theorem **??** shows that the power series expansions for sinh and cosh are convergent for all x.

The following result gives some of the easily determined properties of sinh and cosh.

3.6.22 Proposition (Properties of sinh and cosh) *The functions* sinh *and* cosh *enjoy the following properties:*

(i) $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$ *and* $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$*;*

(ii) sinh *and* cosh *are infinitely differentiable, and furthermore satisfy* $\sinh' = \cosh$ *and* $\cosh' = \sinh$*;*

(iii) $\sinh(-x) = \sinh(x)$ *and* $\cosh(-x) = \cosh(x)$ *for all* $x \in \mathbb{R}$*;*

(iv) $\cosh(x)^2 - \sinh(x)^2 = 1$ *for all* $x \in \mathbb{R}$.

Proof (i) These follows directly from the $\mathbb{R}$-power series definitions for exp, sinh, and cosh.

(ii) This follows from Corollary **??** and the fact that $\mathbb{R}$-convergent power series can be differentiated term-by-term.

(iii) These follow directly from the $\mathbb{R}$-power series for sinh and cosh.

(iv) This can be proved directly using part (i). ■

Also sometimes useful is the ***hyperbolic tangent function*** $\tanh\colon \mathbb{R} \to \mathbb{R}$ defined by $\tanh(x) = \frac{\sinh(x)}{\cosh(x)}$.

Exercises

3.6.1 For representative values of $a \in \mathbb{R}_{>0}$, give the graph of P_a, showing the features outlined in Proposition 3.6.10.

3.6.2 For representative values of $a \in \mathbb{R}$, give the graph of P^a, showing the features outlined in Proposition 3.6.11.

3.6.3 Prove the following trigonometric identities:

(a) $\cos a \cos b = \frac{1}{2}(\cos(a+b) + \cos(a-b))$;

(b) $\cos a \sin b = \frac{1}{2}(\sin(a+b) - \sin(a-b))$;

(c) $\sin a \sin b = \frac{1}{2}(\cos(a-b) - \cos(a+b))$.

3.6.4 Prove the following trigonometric identities:

(a)

3.6.5 Show that tanh is injective.

Chapter 4
Algebraic structures

During the course of these volumes, we shall occasionally, sometimes in essential ways, make use of certain ideas from abstract algebra, particular abstract linear algebra. In this chapter we provide the necessary background in abstract algebra, saving the subject of linear algebra for Chapter **??**. Our idea is to provide sufficient detail to give some context to the instances when we make use of algebra.

Do I need to read this chapter? Provided that the reader is comfortable with the very basic arithmetic ideas concerning integers, real numbers, complex numbers, and polynomials, the material in Sections 4.1–**??** can probably be skipped until it is needed in the course of the text. When it is needed, however, a reader with little exposure to abstract algebra can expect to expend some effort even for the basic material we present here. The material in Section 4.3 appears immediately in Chapter 8 in our initial consideration of the concept of spaces of signals. For this reason, the material should be considered essential. However, it is possible that certain parts of the chapter can be skimmed at a first reading, since the most essential concept is that of a vector space as defined and discussed in Section 4.3. The preparatory material of Sections 4.1–**??** in not essential for understanding what a vector space is, particularly if one is comfortable with the algebraic structure of the set $\mathbb{R}$ of real numbers and the set $\mathbb{C}$ of complex numbers. Section **??** will not be important for significant portions of the text, so can easily be skipped until needed or wanted. •

Contents

4.1 Groups 287
4.1.1 Definitions and basic properties 287
4.1.2 Subgroups 290
4.1.3 Quotients 292
4.1.4 Group homomorphisms 294
4.1.5 The symmetric group 296
4.2 Fields 303
4.2.1 Definitions and basic properties 303
4.2.2 Fraction fields 304
4.2.3 Subfields, field homomorphisms, and characteristic 305
4.3 Vector spaces 309

4.3.1 Definitions and basic properties . . . 309
4.3.2 Subspaces . . . 312
4.3.3 Linear independence . . . 315
4.3.4 Basis and dimension . . . 320
4.3.5 Intersections, sums, and products . . . 327
4.3.6 Complements and quotients . . . 334
4.3.7 Complexification of $\mathbb{R}$-vector spaces . . . 337
4.3.8 Extending the scalars for a vector space . . . 340
4.3.9 Notes . . . 342

Section 4.1

Groups

One of the basic structures in mathematics is that of a group. A group structure often forms the building block for more particular algebraic structures.

Do I need to read this section? Since the material in this section is not difficult, although it is abstract, it may be useful reading for those who feel as if they need to get some familiarity with simple abstract constructions and proofs. The content of the section itself is necessary reading for those who want to understand the material in Sections **??**–**??**. •

4.1.1 Definitions and basic properties

There are a few structures possessing less structure than a group, so we first define these. Many of our definitions of algebraic structure involve the notion of a "binary operation," so let us make this precise.

4.1.1 Definition (Binary operation) A ***binary operation*** on a set S is a map $B\colon S\times S\to S$. A pair (S,B) where B is a binary operation on S is a ***magma***. •

We begin with one of the most basic of algebraic structures, even more basic than a group.

4.1.2 Definition (Semigroup) A ***semigroup*** is a nonempty set S with a binary operation on S, denoted by $(s_1,s_2)\mapsto s_1\cdot s_2$, having the property that

(i) $(s_1\cdot s_2)\cdot s_3 = s_1\cdot(s_2\cdot s_3)$ for all $s_1,s_2,s_3\in\mathsf{S}$ (***associativity***). •

Slightly more structured than a semigroup is the idea of a monoid.

4.1.3 Definition (Monoid) A ***monoid*** is a nonempty set M with a binary operation on M, denoted by $(m_1,m_2)\mapsto m_1\cdot m_2$, having the following properties:

(i) $m_1\cdot(m_2\cdot m_3) = (m_1\cdot m_2)\cdot m_3$ for all $m_1,m_2,m_3\in\mathsf{M}$ (***associativity***);

(ii) there exists $e\in\mathsf{M}$ such that $m\cdot e = e\cdot m = m$ for all $m\in\mathsf{M}$ (***identity element***). •

Now we define what we mean by a group.

4.1.4 Definition (Group) A ***group*** is a nonempty set G endowed with a binary operation, denoted by $(g_1,g_2)\mapsto g_1\cdot g_2$, having the following properties:

(i) $g_1\cdot(g_2\cdot g_3) = (g_1\cdot g_2)\cdot g_3$ for all $g_1,g_2,g_3\in\mathsf{G}$ (***associativity***);

(ii) there exists $e\in\mathsf{G}$ such that $g\cdot e = e\cdot g = g$ for all $g\in\mathsf{G}$ (***identity element***);

(iii) for each $g\in\mathsf{G}$ there exists $g^{-1}\in\mathsf{G}$ such that $g\cdot g^{-1} = g^{-1}\cdot g = e$ (***inverse element***).

A group is ***Abelian*** if $g_1\cdot g_2 = g_2\cdot g_1$ for all $g_1,g_2\in\mathsf{G}$. •

As we did when we defined the operation of multiplication in $\mathbb{R}$, we will often omit the symbol "·" for the binary operation in a group (or semigroup or monoid), and simply write g_1g_2 in place of $g_1 \cdot g_2$. When talking simultaneously about more than one group, it is sometimes advantageous to denote the identity element of a group G by e_{G}.

Clearly the following inclusions hold:

$$\text{Semigroups} \subseteq \text{Monoids} \subseteq \text{Groups}.$$

Throughout these volumes, we shall encounter many examples of groups. For the moment, let us give some very simple examples that illustrate the difference between the ideas of a semigroup, monoid, and group.

4.1.5 Examples (Semigroups, monoids, and groups)

1. A singleton $\{x\}$ with the (only possible) binary operation $x \cdot x = x$ is a group with identity element x and with inverse element defined by $x^{-1} = x$.
2. The set $\mathbb{Z}_{>0}$ with the binary operation of addition is a semigroup. However, it is not a monoid since it has no identity element, and it is not a group, because it has no identity element and so there are also no inverse elements.
3. The set $\mathbb{Z}_{>0}$ with the binary operation of multiplication is a monoid with identity element $e = 1$. It is not a group.
4. The set $\mathbb{Z}_{\geq 0}$ with the binary operation of addition is a monoid with identity element 0, but not a group.
5. The set $\mathbb{Z}_{\geq 0}$ with the binary operation of multiplication is a monoid with identity element 1. It is not a group.
6. The set $\mathbb{Z}$ with the binary operation of addition is a group with identity element 0, and with inverse defined by $k^{-1} = -k$.
7. The set $\mathbb{Z}$ with the binary operation of multiplication is a monoid with identity 1, but it is not a group.
8. The sets $\mathbb{Q}$ and $\mathbb{R}$ with the binary operations of addition are groups with identity element 0 and with inverse defined by $x^{-1} = -x$.
9. The sets $\mathbb{Q}$ and $\mathbb{R}$ with the binary operations of multiplication are monoids with identity element 1. They are not groups.
10. The sets $\mathbb{Q}^* \triangleq \mathbb{Q} \setminus \{0\}$ and $\mathbb{R}^* \triangleq \mathbb{R} \setminus \{0\}$ with the binary operation of multiplication are groups with identity element 1 and with inverse given by $x^{-1} = \frac{1}{x}$.
11. Let $\mathfrak{S}_k$, $k \in \mathbb{Z}_{>0}$, denote the set of bijections of the set $\{1, \dots, k\}$, and equip $\mathfrak{S}_k$ with the binary operation $(\sigma_1, \sigma_2) \mapsto \sigma_1 \circ \sigma_2$. One can easily verify that $\mathfrak{S}_k$ is a group with identity given by the identity map, and with inverse given by the inverse map. This group is called the ***permutation group*** or the ***symmetric group*** on k symbols. It is conventional to represent a permutation $\sigma \in \mathfrak{S}_k$ using the following matrix-type representation:

$$\begin{pmatrix} 1 & 2 & \cdots & k \\ \sigma(1) & \sigma(2) & \cdots & \sigma(k) \end{pmatrix}.$$

Thus the first row contains the elements $\{1,\dots,k\}$ in order, and the second row contains the images of these elements under σ.

We claim that $\mathfrak{S}_k$ is Abelian when $k \in \{1,2\}$, and otherwise is not Abelian. We leave it to the reader to check directly that $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are Abelian. Let us show that $\mathfrak{S}_3$ is not Abelian. Define $\sigma_1, \sigma_2 \in \mathfrak{S}_3$ by

$$\begin{aligned} &\sigma_1(1) = 2, \quad \sigma_1(2) = 1, \quad \sigma_1(3) = 3,\\ &\sigma_2(1) = 1, \quad \sigma_2(2) = 3, \quad \sigma_2(3) = 2. \end{aligned}$$

One can then verify that

$$\begin{aligned} &\sigma_1 \circ \sigma_2(1) = 2, \quad \sigma_1 \circ \sigma_2(2) = 3, \quad \sigma_1 \circ \sigma_2(3) = 1,\\ &\sigma_2 \circ \sigma_1(1) = 3, \quad \sigma_2 \circ \sigma_1(2) = 1, \quad \sigma_2 \circ \sigma_1(3) = 2. \end{aligned}$$

Thus $\mathfrak{S}_3$ in indeed not Abelian.

That $\mathfrak{S}_k$ is not Abelian for $k > 3$ follows since in Example 4.1.12–7 we will show that $\mathfrak{S}_3$ is a isomorphic to a subgroup of $\mathfrak{S}_k$ (asking the readers forgiveness that the terms "isomorphic" and "subgroup" have yet to be defined; they will be shortly).

We shall have more to say about the symmetric group in Section 4.1.5.

All groups in the above list may be verified to be Abelian, with the exception of the permutation group on k symbols for $k \geq 2$. •

Having introduced the notions of a semigroup and monoid, we shall not make much use of them. They are, however, useful in illustrating what a group is and is not.

The following properties of groups are more or less easily verified, and we leave the verifications to the reader as Exercise 4.1.1.

4.1.6 Proposition (Elementary properties of groups) *If* G *is a group, then the following statements hold:*

(i) there is exactly one element $\mathrm{e} \in \mathsf{G}$ *that satisfies* $\mathrm{g} \cdot \mathrm{e} = \mathrm{e} \cdot \mathrm{g} = \mathrm{g}$ *for all* $\mathrm{g} \in \mathsf{G}$*, i.e., the identity element in a group is unique;*

(ii) for $\mathrm{g} \in \mathsf{G}$*, there exists exactly one element* $\mathrm{g}' \in \mathsf{G}$ *such that* $\mathrm{g}' \cdot \mathrm{g} = \mathrm{g} \cdot \mathrm{g}' = \mathrm{e}$*, i.e., inverse elements are unique;*

(iii) for $\mathrm{g} \in \mathsf{G}$*,* $(\mathrm{g}^{-1})^{-1} = \mathrm{g}$*;*

(iv) for $\mathrm{g}_1, \mathrm{g}_2 \in \mathsf{G}$*,* $(\mathrm{g}_1 \cdot \mathrm{g}_2)^{-1} = \mathrm{g}_2^{-1} \cdot \mathrm{g}_1^{-1}$*;*

(v) if $\mathrm{g}_1, \mathrm{g}_2, \mathrm{h} \in \mathsf{G}$ *satisfy* $\mathrm{h} \cdot \mathrm{g}_1 = \mathrm{h} \cdot \mathrm{g}_2$*, then* $\mathrm{g}_1 = \mathrm{g}_2$*;*

(vi) if $\mathrm{g}_1, \mathrm{g}_2, \mathrm{h} \in \mathsf{G}$ *satisfy* $\mathrm{g}_1 \cdot \mathrm{h} = \mathrm{g}_2 \cdot \mathrm{h}$*, then* $\mathrm{g}_1 = \mathrm{g}_2$*;*

(vii) if $\mathrm{g}_1, \mathrm{g}_2 \in \mathsf{G}$*, then there exists a unique* $\mathrm{h} \in \mathsf{G}$ *such that* $\mathrm{g}_1 \cdot \mathrm{h} = \mathrm{g}_2$*;*

(viii) if $\mathrm{g}_1, \mathrm{g}_2 \in \mathsf{G}$*, then there exists a unique* $\mathrm{h} \in \mathsf{G}$ *such that* $\mathrm{h} \cdot \mathrm{g}_1 = \mathrm{g}_2$*.*

There is some useful notation associated with iterated group multiplication. Namely, if G is a semigroup, if $g \in \mathsf{G}$, and if $k \in \mathbb{Z}_{>0}$, then we define $g^k \in \mathsf{G}$ iteratively by $g^1 = g$ and $g^k = g \cdot g^{k-1}$. The following result records the fact that this notation behaves as we expect.

4.1.7 Proposition (Properties of g^k) *If* G *is a semigroup, if* $g \in G$*, and if* $k_1, k_2 \in \mathbb{Z}_{>0}$*, then the following statements hold:*

(i) $g^{k_1} \cdot g^{k_2} = g^{k_1+k_2}$*;*

(ii) $(g^{k_1})^{k_2} = g^{k_1 k_2}$*.*

Proof (i) Let $g \in G$ and $k_1 \in \mathbb{Z}_{>0}$. If $k_2 = 1$ then, by definition,

$$g^{k_1} \cdot g^{k_2} = g^{k_1} \cdot g = g^{k_1+1} = g^{k_1+k_2},$$

so the result holds for $k_2 = 1$. Now suppose that the result holds for $k_2 \in \{1, \dots, k\}$. Then, if $k_2 = k + 1$,

$$g^{k_1} g^{k_2} = g^{k_1} \cdot g^{k+1} = g^{k_1} \cdot g^k \cdot g = g^{k_1+k} \cdot g = g^{k_1+k+1} = g^{k_1+k_2},$$

giving the result by induction on k_2.

(ii) Let $g \in G$ and $k_1 \in \mathbb{Z}_{>0}$. If $k_2 = 1$ then clearly $(g^{k_1})^{k_2} = g^{k_1 k_2}$. Now suppose that the result holds for $k_2 \in \{1, \dots, k\}$, and for $k_2 = k + 1$ compute

$$(g^{k_1})^{k_2} = (g^{k_1})^{k+1} = (g^{k_1})^k \cdot g^{k_1} = g^{k_1 k} \cdot g^{k_1} = g^{k_1 k + k_1} = g^{k_1(k+1)} = g^{k_1 k_2},$$

giving the result by induction on k_2. ■

4.1.8 Notation (g^k for Abelian groups) When a group is Abelian, then the group operation is sometimes thought of as addition, since it shares the property of commutativity possessed by addition. In such cases, one often write "kg" in place of "g^k" to reflect the idea that the group operation is "additive." •

4.1.2 Subgroups

It is often useful to consider subsets of groups that respect the group operation.

4.1.9 Definition (Subgroup) A nonempty subset H of a group G is a ***subgroup*** if

(i) $h_1 \cdot h_2 \in H$ for all $h_1, h_2 \in H$ and

(ii) $h^{-1} \in H$ for all $h \in H$. •

The following property of subgroups are easily verified, as the reader can see by doing Exercise 4.1.5.

4.1.10 Proposition (A subgroup is a group) *A nonempty subset* $H \subseteq G$ *of a group* G *is a subgroup if and only if* H *is a group using the binary operation of multiplication in* G*, restricted to* H*.*

4.1.11 Remark (On sub"objects") Mathematics can be perhaps thought of as the study of sets having some prescribed structure. It is frequent that one is interested in subsets which inherit this structure from the superset. Such subsets are almost always named with the prefix "sub." The above notion of a subgroup is our first encounter with this idea, although it will come up frequently in these volumes. •

Let us give some examples of subgroups.

4.1.12 Examples (Subgroups)

1. For any group G, $\{e\}$ is a subgroup, often called the ***trivial subgroup***.
2. Let $k \in \mathbb{Z}_{>0}$. The subset $k\mathbb{Z}$ of $\mathbb{Z}$ defined by
$$k\mathbb{Z} = \{kj \mid j \in \mathbb{Z}\}$$
(i.e., $k\mathbb{Z}$ consists of multiples of k) is a subgroup of $\mathbb{Z}$ if $\mathbb{Z}$ possesses the binary operation of addition.
3. $\mathbb{Z}$ and $\mathbb{Q}$ are subgroups of $\mathbb{R}$ if $\mathbb{R}$ possesses the binary operation of addition.
4. $\mathbb{Q}^*$ is a subgroup of $\mathbb{R}^*$ if $\mathbb{R}$ possesses the binary operation of multiplication.
5. $\mathbb{Z}$ is not a subgroup of $\mathbb{Q}$ if $\mathbb{Q}$ possesses the binary operation of multiplication.
6. Neither $\mathbb{Z}_{>0}$ nor $\mathbb{Z}_{\geq 0}$ are subgroups of $\mathbb{Z}$ if $\mathbb{Z}$ possesses the binary operation of addition.
7. Let $l, k \in \mathbb{Z}_{>0}$ with $l < k$. Let $\mathfrak{S}_{l,k}$ be the subset of $\mathfrak{S}_k$ defined by
$$\mathfrak{S}_{l,k} = \{\sigma \in \mathfrak{S}_k \mid \sigma(j) = j, \ j > l\}.$$
We claim that $\mathfrak{S}_{l,k}$ is a subgroup of $\mathfrak{S}_k$. It is clear by definition that, if $\sigma_1, \sigma_2 \in \mathfrak{S}_{l,k}$, then $h_1 \circ h_2 \in \mathfrak{S}_{l,k}$. If $\sigma \in \mathfrak{S}_{l,k}$ then let us write $\psi(j) = \sigma(j)$ for $j \in \{1, \dots, l\}$. This then defines $\psi \in \mathfrak{S}_l$. One can then directly verify that σ^{-1} is defined by
$$\sigma^{-1}(j) = \begin{cases} \psi^{-1}(j), & j \in \{1, \dots, l\}, \\ j, & j > l. \end{cases}$$
Thus $\sigma^{-1} \in \mathfrak{S}_{l,k}$, as desired.
Note that our above computations show that essentially $\mathfrak{S}_{l,k}$ consists of a copy of $\mathfrak{S}_l$ sitting inside $\mathfrak{S}_k$. In the language we are about to introduce in Definition 4.1.20, $\mathfrak{S}_{l,k}$ is isomorphic to $\mathfrak{S}_l$ (see Example 4.1.23–2). •

An important idea in many algebraic settings is that of the smallest subobject containing some subset. For groups this construction rests on the following result.

4.1.13 Proposition (Existence of subgroup generated by a subset) *Let G be a group and let $\mathsf{S} \subseteq \mathsf{G}$. Then there exists a subgroup $\mathsf{H}_\mathsf{S} \subseteq \mathsf{G}$ such that*

(i) $\mathsf{S} \subseteq \mathsf{H}_\mathsf{S}$ *and*

(ii) if $\mathsf{H} \subseteq \mathsf{G}$ *is a subgroup for which* $\mathsf{S} \subseteq \mathsf{H}$ *then* $\mathsf{H}_\mathsf{S} \subseteq \mathsf{H}$.

Moreover,
$$\mathsf{H}_\mathsf{S} = \{\mathsf{g}_1 \cdots \mathsf{g}_\mathsf{k} \mid \mathsf{k} \in \mathbb{Z}_{>0}, \ \mathsf{g}_\mathsf{j} \in \mathsf{S} \textit{ or } \mathsf{g}_\mathsf{j}^{-1} \in \mathsf{S}, \ \mathsf{j} \in \{1, \dots, \mathsf{k}\}\}$$
is the unique subgroup having the above two properties.

Proof Let
$$\mathscr{H}_S = \{\mathsf{H} \subseteq \mathsf{G} \mid \mathsf{H} \text{ is a subgroup with } S \subseteq \mathsf{H}\}.$$
Since $\mathsf{G} \in \mathscr{H}_S$ it follows that $\mathscr{H}_S$ is nonempty. We claim that $\mathsf{H}_S \triangleq \cap_{\mathsf{H} \in \mathscr{H}_S} \mathsf{H}$ has the required properties. First let $g \in S$. Then $g \in \mathsf{H}$ for every $\mathsf{H} \in \mathscr{H}_S$. Thus $g \in \mathsf{H}_S$ and so $S \subseteq \mathsf{H}_S$. Now let $g_1, g_2 \in \mathsf{H}_S$. Then $g_1, g_2 \in \mathsf{H}$ for every $\mathsf{H} \in \mathscr{H}_S$ and so $g_1 \cdot g_2 \in \mathsf{H}$

for every $\mathsf{H} \in \mathscr{H}_S$. Similarly, if $g \in \mathsf{H}$ for every $\mathsf{H} \in \mathscr{H}_S$ then $g^{-1} \in \mathsf{H}$ for every $\mathsf{H} \in \mathscr{H}_S$. Thus H_S is a subgroup containing S. Furthermore, if H is a subgroup containing S and if $g \in \mathsf{H}_S$ then clearly $g \in \mathsf{H}$ since $\mathsf{H} \in \mathscr{H}_S$. Thus $\mathsf{H}_S \subseteq \mathsf{H}$. We, moreover, claim that there is only one subgroup having the two stated properties. Indeed, suppose that $\mathsf{H}'_S \subseteq \mathsf{G}$ is a subgroup containing S and if H'_S is contained in any subgroup containing S. Then $\mathsf{H}'_S \subseteq \mathsf{H}_S$. Moreover, since $\mathsf{H}'_S \in \mathscr{H}_S$ we have $\mathsf{H}_S \subseteq \mathsf{H}'_S$. Thus $\mathsf{H}'_S = \mathsf{H}_S$.

To prove the final assertion it now suffices to show that

$$\mathsf{H}'_S = \{g_1 \cdots g_k \mid k \in \mathbb{Z}_{>0},\ g_j \in S \text{ or } g_j^{-1} \in S,\ j \in \{1, \dots, k\}\}$$

is a subgroup containing S and has the property that $\mathsf{H}'_S \subseteq \mathsf{H}$ for any subgroup H containing S. Clearly $S \subseteq \mathsf{H}'_S$. Now let

$$g_1 \cdots g_k, g'_1, \dots, g'_{k'} \in \mathsf{H}'_S.$$

Then clearly

$$g_1 \cdots g_k \cdot g'_1, \dots, g'_{k'} \in \mathsf{H}'_S.$$

Moreover,

$$(g_1 \cdots g_k)^{-1} = g_k^{-1} \cdots g_1^{-1} \in \mathsf{H}'_S$$

and so H'_S is a subgroup. Now let H be a subgroup containing S. Then $g_1 \cdot g_2 \in \mathsf{H}$ and $g^{-1} \in \mathsf{H}$ for every $g, g_1, g_2 \in S$. This means that $g_1 \cdots g_k \in \mathsf{H}$ for every $g_1, \dots, g_k \in \mathsf{G}$ such that either g_j or g_j^{-1} are in S, $j \in \{1, \dots, k\}$. Thus $\mathsf{H}'_S \subseteq \mathsf{H}$ and so we conclude that $\mathsf{H}'_S = \mathsf{H}_S$. ■

4.1.14 Definition (Subgroup generated by a subset) If G is a group and if $S \subseteq \mathsf{G}$, the subgroup H_S of Proposition 4.1.13 is the ***subgroup generated by* S**. •

4.1.3 Quotients

Let us now turn to some important ideas connected with subgroups.

4.1.15 Definition (Left and right cosets) Let G be a group with H a subgroup.

(i) The ***left coset*** of H through $g \in \mathsf{G}$ is the set $g\mathsf{H} = \{gh \mid h \in \mathsf{H}\}$.

(ii) The ***right coset*** of H through $g \in \mathsf{G}$ is the set $\mathsf{H}g = \{hg \mid h \in \mathsf{H}\}$.

The set of left (resp. right) cosets is denoted by G/H (resp. $\mathsf{H}\backslash\mathsf{G}$), and the map assigning to $g \in \mathsf{G}$ the coset $g\mathsf{H} \in \mathsf{G}/\mathsf{H}$ (resp. $\mathsf{H}g \in \mathsf{H}\backslash\mathsf{G}$) is denoted by π_{H} (resp. ${}_{\mathsf{H}}\pi$), and is called the ***canonical projection***. •

Of course, if G is Abelian, then $g\mathsf{H} = \mathsf{H}g$ for each $g \in \mathsf{G}$, and, as a consequence, the sets G/H and $\mathsf{H}\backslash\mathsf{G}$ are the same. It is common to refer to G/H or $\mathsf{H}\backslash\mathsf{G}$ as the ***quotient*** of G by H.

An alternative description of cosets is given by the following result.

4.1.16 Proposition (Cosets as equivalence classes) *The set* G/H *(resp.* $\mathsf{H}\backslash\mathsf{G}$*) is the same as the set of equivalence classes in* G *associated to the equivalence relation* $\mathsf{g}_1 \sim \mathsf{g}_2$ *if* $\mathsf{g}_2^{-1}\mathsf{g}_1 \in \mathsf{H}$ *(resp.* $\mathsf{g}_2\mathsf{g}_1^{-1} \in \mathsf{H}$*).*

Proof We prove the proposition only for left cosets, and the proof for right cosets follows, *mutatis mutandis*. First let us prove that the relation defined by $g_1 \sim g_2$ if $g_2^{-1}g_1 \in \mathsf{H}$ is an equivalence relation.

1. Note that $g^{-1}g = e \in \mathsf{H}$, so the relation is reflexive.
2. If $g_1 \sim g_2$ then $g_2^{-1}g_1 \in \mathsf{H}$, which implies that $(g_2^{-1}g_1)^{-1} \in \mathsf{H}$ since H is a subgroup. By Proposition 4.1.6 this means that $g_1^{-1}g_2 \in \mathsf{H}$; i.e., that $g_2 \sim g_1$. Thus the relation is symmetric.
3. If $g_1 \sim g_2$ and $g_2 \sim g_3$, or equivalently that $g_2 \sim g_1$ and $g_3 \sim g_2$, then $g_1^{-1}g_2, g_2^{-1}g_3 \in \mathsf{H}$. Then, since H is a subgroup,
$$(g_1^{-1}g_2)(g_2^{-1}g_3) \in \mathsf{H} \quad \Longrightarrow \quad g_1^{-1}g_3 \in \mathsf{H}.$$
Thus $g_3 \sim g_1$, or $g_1 \sim g_3$, and the relation is transitive.

Now let $g \in \mathsf{G}$ and let $g' \in g\mathsf{H}$. Then $g' = gh$ for some $h \in \mathsf{H}$, so $g^{-1}g' \in \mathsf{H}$, so $g' \sim g$. Conversely, suppose that $g' \sim g$ so that $g^{-1}g' = h$ for some $h \in \mathsf{H}$. Then $g' = gh$, so $g' \in g\mathsf{H}$. This gives the result. ■

Let us give some examples of cosets and collections of cosets.

4.1.17 Examples (Cosets)

1. Let $k \in \mathbb{Z}_{>0}$. Consider the group $\mathbb{Z}$ with the binary operation of addition, and also consider the subgroup $k\mathbb{Z}$ consisting of multiples of k. We claim that $\mathbb{Z}/k\mathbb{Z}$ is a set with k elements. Using the Theorem **??** below, we see that every element of $\mathbb{Z}$ lies in the coset of exactly one of the elements from the set $\{0, 1, \dots, k-1\}$, which gives our claim. For reasons which will become clear in Example **??**–**??** it is convenient to denote the coset through $j \in \mathbb{Z}$ by $j + k\mathbb{Z}$. We will frequently encounter the group $\mathbb{Z}/k\mathbb{Z}$, and so give it the shorthand $\mathbb{Z}_k$.
2. Consider the group $\mathbb{R}$ equipped with the binary operation of addition, and consider the subgroup $\mathbb{Q}$. We claim that the set $\mathbb{R}/\mathbb{Q}$ is uncountable. Indeed, if it were not, then this would imply that $\mathbb{R}$ is the countable union of cosets, and each coset itself must be countable. That is to say, if $\mathbb{R}/\mathbb{Q}$ is countable, then $\mathbb{R}$ is a countable union of countable sets. But, by Proposition **??**, this means that $\mathbb{R}$ is countable. However, in Exercise 2.1.4 the reader is asked to show $\mathbb{R}$ is actually not countable. The contradiction proves that $\mathbb{R}/\mathbb{Q}$ is uncountable. Further investigation of $\mathbb{R}/\mathbb{Q}$ takes one into the topic of field extensions, which we consider very briefly in Section 4.2.3, and then into Galois theory, which is somewhat beyond our focus here.
3. Consider the permutation group $\mathfrak{S}_3$ in 3 symbols and consider the subgroup $\mathfrak{S}_{2,3}$, which is isomorphic to $\mathfrak{S}_2$ as we showed in Example 4.1.23–2. Let us describe the cosets of $\mathfrak{S}_3/\mathfrak{S}_{2,3}$. Suppose that $\sigma_1, \sigma_2 \in \mathfrak{S}_3$ lie in the same coset of $\mathfrak{S}_{2,3}$. Then it must hold that $\sigma_1 \circ \sigma_2^{-1}(3) = 3$, or equivalently that $\sigma_1^{-1}(3) = \sigma_2^{-1}(3)$. Thus cosets are identified by their having in common the fact that the same

elements in $\{1,2,3\}$ are images of the element 3. The cosets then are easily seen to be

(a) $\left\{\begin{pmatrix}1 & 2 & 3\\ 1 & 2 & 3\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{pmatrix}\right\}$,

(b) $\left\{\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\\ 3 & 2 & 1\end{pmatrix}\right\}$, and

(c) $\left\{\begin{pmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix}\right\}$. •

Next we discuss a particular sort of subgroup that, as we shall see, is distinguished by the structure of its set of cosets.

4.1.18 Definition (Normal subgroup) A subgroup H of a group G is a ***normal subgroup*** if $g\mathsf{H} = \mathsf{H}g$ for all $g \in \mathsf{G}$. •

The following result explains why normal subgroups are interesting.

4.1.19 Proposition (Quotients by normal subgroups are groups) *Let* N *be a normal subgroup of* G *and define a binary operation on* G/N *by*

$$(\mathrm{g}_1\mathsf{N}, \mathrm{g}_2\mathsf{N}) \mapsto (\mathrm{g}_1\mathrm{g}_2)\mathsf{N}.$$

Then this binary operation satisfies the conditions for group multiplication.

Proof First let us show that this binary operation is well-defined. Let $g_1, g_2, h_1, h_2 \in \mathsf{G}$ satisfy $g_1\mathsf{N} = h_1\mathsf{N}$ and $g_2\mathsf{N} = h_2\mathsf{N}$. Then we must have $g_1^{-1}h_1 = n_1$ and $g_2^{-1}h_2 = n_2$ for $n_1, n_2 \in \mathsf{N}$, and then we compute

$$\begin{aligned}(h_1h_2\mathsf{N}) &= \{h_1h_2n \mid n \in \mathsf{N}\} = \{g_1n_1g_2n_2n \mid n \in \mathsf{N}\}\\ &= \{g_1g_2n_3n_2n \mid n \in \mathsf{N}\} = \{g_1g_2n \mid n \in \mathsf{N}\} = (g_1g_2)\mathsf{N},\end{aligned}$$

where $n_3 \in \mathsf{N}$ is defined so that $n_1g_2 = g_2n_3$, this being possible by Exercise 4.1.8 since N is normal.

To then verify that the (now) well-defined binary operation satisfies the conditions for group multiplication is trivial. ■

4.1.4 Group homomorphisms

Another important concept for groups, and for many other structures in mathematics, is that of a map that preserves the structure.

4.1.20 Definition (Group homomorphism, epimorphism, monomorphism, and isomorphism) For semigroups (resp. monoids, groups) G and H, a map $\phi\colon \mathsf{G} \to \mathsf{H}$ is a:

(i) ***semigroup*** (resp. ***monoid***, ***group***) ***homomorphism***, or simply a ***homomorphism***, if $\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2)$ for all $g_1, g_2 \in \mathsf{G}$;

(ii) ***epimorphism*** if it is a surjective homomorphism;

(iii) ***monomorphism*** if it is an injective homomorphism;

(iv) ***isomorphism*** if it a bijective homomorphism. •

We shall mainly be concerned with group homomorphisms, although homomorphisms of semigroups and monoids will arise at times.

4.1.21 Remark (On morphisms of various sorts) As with the idea of a sub"object" as discussed in Remark 4.1.11, the idea of a map between sets that preserves the structure of those sets, e.g., the group structure in the case of a group homomorphism, is of fundamental importance. The expression "morphosis" comes from Greek for "form," whereas the prefixes "homo," "epi," "mono," and "isos" are from the Greek for roughly "alike," "on," "one," and "equal," respectively. •

The following result gives a couple of basic properties of homomorphisms.

4.1.22 Proposition (Properties of group homomorphisms) *If* G *and* H *are monoids and if* $\phi\colon \mathsf{G} \to \mathsf{H}$ *is a monoid homomorphism, then*

(i) $\phi(\mathrm{e}_{\mathsf{G}}) = \mathrm{e}_{\mathsf{H}}$*, and*

(ii) if G *and* H *are additionally groups, then* $\phi(\mathrm{g}^{-1}) = (\phi(\mathrm{g}))^{-1}$*.*

Proof (i) Let $g \in \mathsf{G}$ and note that

$$\phi(e_{\mathsf{G}}g) = \phi(ge_{\mathsf{G}}) = \phi(e_{\mathsf{G}})\phi(g) = \phi(g)\phi(e_{\mathsf{G}}) = \phi(g).$$

In particular, $\phi(g)\phi(e_{\mathsf{G}}) = \phi(g)e_{\mathsf{H}}$, and the result follows by multiplication by $\phi(g)^{-1}$.

(ii) Now, if $g \in \mathsf{G}$ then $\phi(g)\phi(g^{-1}) = \phi(gg^{-1}) = \phi(e_{\mathsf{G}}) = e_{\mathsf{H}}$, which shows that $\phi(g^{-1}) = (\phi(g))^{-1}$. ■

4.1.23 Examples (Group homomorphisms)

1. If G and H are groups with identity elements e_{G} and e_{H}, respectively, then the map $\phi\colon \mathsf{G} \to \mathsf{H}$ defined by $\phi(g) = e_{\mathsf{H}}$ for all $g \in \mathsf{G}$ is readily verified to be a homomorphism. It is an epimorphism if and only if $\mathsf{H} = \{e_{\mathsf{H}}\}$ and a monomorphism if and only if $\mathsf{G} = \{e_{\mathsf{G}}\}$.
2. Let $l, k \in \mathbb{Z}_{>0}$ with $l < k$. The map $\phi\colon \mathfrak{S}_l \to \mathfrak{S}_k$ defined by

$$\phi(\sigma)(j) = \begin{cases} \sigma(j), & j \in \{1, \dots, l\}, \\ j, & j > l \end{cases}$$

is verified to be a monomorphism. In fact, it is easily verified to be an isomorphism from $\mathfrak{S}_l$ to $\mathfrak{S}_{l,k} \subseteq \mathfrak{S}_k$. •

Associated to every homomorphism of groups are two important subsets, one of the domain and one of the codomain of the homomorphism.

4.1.24 Definition (Image and kernel of group homomorphism) Let G and H be groups and let $\phi\colon \mathsf{G} \to \mathsf{H}$ be a homomorphism.

(i) The ***image*** of ϕ is $\mathrm{image}(\phi) = \{\phi(g) \mid g \in \mathsf{G}\}$.

(ii) The ***kernel*** of ϕ is $\ker(\phi) = \{g \in \mathsf{G} \mid \phi(g) = e_{\mathsf{H}}\}$. •

The image and the kernel have useful properties relative to the group structure.

4.1.25 Proposition (Image and kernel are subgroups) *If* G *and* H *are groups and if* $\phi\colon \mathsf{G} \to \mathsf{H}$ *is a homomorphism, then*

(i) $\mathrm{image}(\phi)$ *is a subgroup of* H *and*

(ii) $\ker(\phi)$ *is a normal subgroup of* G.

Proof (i) If $g_1, g_2 \in \mathsf{G}$ then $\phi(g_1)\phi(g_2) = \phi(g_1g_2) \in \mathrm{image}(\phi)$. From part (ii) of Proposition 4.1.22 we have $(\phi(g))^{-1} \in \mathrm{image}(\phi)$ for every $g \in \mathsf{G}$.

(ii) Let $g_1, g_2 \in \ker(\phi)$. Then $\phi(g_1g_2) = \phi(g_1)\phi(g_2) = e_{\mathsf{H}}$ so that $g_1g_2 \in \ker(\phi)$. If $g \in \ker(\phi)$ then

$$e_{\mathsf{H}} = \phi(e_{\mathsf{G}}) = \phi(gg^{-1}) = \phi(g)\phi(g^{-1}) = \phi(g^{-1}).$$

Thus $g^{-1} \in \ker(\phi)$, and so $\ker(\phi)$ is a subgroup. To show that $\ker(\phi)$ is normal, let $g \in \mathsf{G}$ and let $h \in \ker(\phi)$. Then

$$\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g^{-1}) = \phi(g)\phi(g^{-1}) = e_{\mathsf{H}}.$$

Thus $ghg^{-1} \in \ker(\phi)$ for every $g \in \mathsf{G}$ and $h \in \ker(\phi)$. The result now follows by Exercise 4.1.8. ■

The following result characterising group monomorphisms is simple, but is one that we use continually, so it is worth recording.

4.1.26 Proposition (Characterisation of monomorphisms) *A group homomorphism* $\phi\colon \mathsf{G} \to \mathsf{H}$ *is a monomorphism if and only if* $\ker(\phi) = \mathrm{e}_{\mathsf{G}}$.

Proof Suppose that $\ker(\phi) = \{e_{\mathsf{G}}\}$ and that $\phi(g_1) = \phi(g_2)$. Then

$$e_{\mathsf{H}} = \phi(g_1)(\phi(g_2))^{-1} = \phi(g_1)\phi(g_2^{-1}) = \phi(g_1g_2^{-1}),$$

implying that $g_1g_2^{-1} \in \ker(\phi)$ whence $g_1 = g_2$, and so ϕ is injective.

Conversely, suppose ϕ is a monomorphism and let $g \in \ker(\phi)$. Thus $\phi(g) = e_{\mathsf{H}}$. However, since ϕ is a monomorphism and since $\phi(e_{\mathsf{G}}) = e_{\mathsf{H}}$, we must have $g = e_{\mathsf{G}}$. ■

4.1.5 The symmetric group

In Example 4.1.5–11 we introduced the symmetric group. We shall have occasion to use some of the structure of the symmetric group, and in this section we collect the pertinent facts.

First of all let us define a simple collection of elements of the symmetric group and some notions associated with them.

4.1.27 Definition (Cycle, transposition, even permutation, odd permutation) Let $k \in \mathbb{Z}_{>0}$.

(i) An element $\sigma \in \mathfrak{S}_k$ is a ***cycle*** if there exists distinct $j_1, \dots, j_m \in \{1, \dots, k\}$ such that

$$\sigma(j_1) = j_2,\ \sigma(j_2) = j_3,\ \cdots,\ \sigma(j_{m-1}) = j_m,\ \sigma(j_m) = j_1,$$

and such that $\sigma(j) = j$ for $j \notin \{j_1, \dots, j_m\}$. The number m is the ***length*** of the cycle. We denote the above cycle by $(j_1\ j_2\ \cdots\ j_m)$.

(ii) An element $\sigma \in \mathfrak{S}_k$ is a ***transposition*** if it is a cycle of length 2. Thus $\sigma = (j_1\ j_2)$ for distinct $j_1, j_2 \in \{1, \dots, k\}$.

(iii) An element $\sigma \in \mathfrak{S}_k$ is ***even*** (resp. ***odd***) if it is a finite product of an even (resp. odd) number of transpositions. •

Let us illustrate the notion of a cycle with an elementary example.

4.1.28 Example (Cycle) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 3 & 2 & 4 \end{pmatrix}$$

is a cycle using the elements 2, 4, and 5, and is written as (2 5 4), representing the fact that $\sigma(2) = 5$, $\sigma(5) = 4$, and $\sigma(4) = 2$. It is clear that one could also write the cycle as (5 4 2) or (4 2 5), and, therefore, the notation we use to represent a cycle is not unique. •

It turns out that every permutation is a product of cycles. If we ask that the cycles have an additional property, then the product is unique. This property is the following.

4.1.29 Definition (Disjoint permutations) Let $k \in \mathbb{Z}_{>0}$. Permutations $\sigma_1, \sigma_2 \in \mathfrak{S}_k$ are ***disjoint*** if, for every $j \in \{1, \dots, k\}$, $\sigma_1(j) \neq j$ implies that $\sigma_2(j) = j$ and $\sigma_2(j) \neq j$ implies that $\sigma_1(j) = j$. •

The idea is that the set of elements of $\{1, \dots, k\}$ not fixed by disjoint permutations are distinct. It is easy to show that disjoint permutations commute; this is Exercise 4.1.12.

We now have the following important structural result describing a typical permutation.

4.1.30 Theorem (Permutations are products of cycles) *Let* $\mathrm{k} \in \mathbb{Z}_{>0}$. *If* $\sigma \in \mathfrak{S}_\mathrm{k}$ *then there exist disjoint cycles* $\sigma_1, \dots, \sigma_\mathrm{r} \in \mathfrak{S}_\mathrm{k}$ *such that* $\sigma = \sigma_1 \circ \cdots \circ \sigma_\mathrm{r}$. *Moreover, if* $\sigma'_1, \dots, \sigma'_{\mathrm{r}'} \in \mathfrak{S}_\mathrm{k}$ *are disjoint permutations such that* $\sigma'_1 \circ \cdots \circ \sigma'_{\mathrm{r}'}$, *then* $\mathrm{r} = \mathrm{r}'$ *and there exists a bijection* $\phi\colon \{1, \dots, \mathrm{r}\} \to \{1, \dots, \mathrm{r}\}$ *such that* $\sigma'_\mathrm{j} = \sigma_{\phi(\mathrm{j})}$, $\mathrm{j} \in \{1, \dots, \mathrm{r}\}$.

Proof For $\sigma \in \mathfrak{S}_k$ and $j \in \{1, \dots, k\}$ let us denote

$$O(\sigma, j) = \{\sigma^m(j) \mid m \in \mathbb{Z}_{\geq 0}\}$$

and suppose that $\operatorname{card}(O(\sigma, j)) = N_{\sigma,j}$.

1 Lemma *With the above notation the following statements hold:*

(i) $\mathrm{j}, \sigma(\mathrm{j}), \dots, \sigma^{\mathrm{N}_{\sigma,\mathrm{j}}-1}(\mathrm{j})$ *are distinct;*

(ii) $\sigma^{\mathrm{N}_{\sigma,\mathrm{j}}}(\mathrm{j}') = \mathrm{j}'$ *for each* $\mathrm{j}' \in \mathrm{O}(\sigma, \mathrm{j})$;

(iii) $\mathrm{O}(\sigma, \mathrm{j}) = \{\mathrm{j}, \sigma(\mathrm{j}), \dots, \sigma^{\mathrm{N}_{\sigma,\mathrm{j}}-1}(\mathrm{j})\}$;

(iv) $\mathrm{O}(\sigma, \mathrm{j}') = \mathrm{O}(\sigma, \mathrm{j})$ *for every* $\mathrm{j}' \in \mathrm{O}(\sigma, \mathrm{j})$.

Proof (i) Suppose that $\sigma^{m_1}(j) = \sigma^{m_2}(j)$ for distinct $m_1, m_2 \in \{0, 1, \dots, N_{\sigma,j} - 1\}$. Suppose that $m_2 > m_1$ so that $\sigma^{m_2 - m_1}(j) = j$ with $m_2 - m_1 \in \{1, \dots, N_{\sigma,j} - 1\}$. For $m \in \mathbb{Z}_{>0}$

let us use the division algorithm for $\mathbb{Z}$ (Theorem **??**) to write $m = q(m_2 - m_1) + r$ for $r \in \{0, 1, \dots, m_2 - m_1 - 1\}$. Then $\sigma^m(j) = \sigma^r(j)$ and so it follows that

$$O(\sigma, j) \subseteq \{j, \sigma(j), \dots, \sigma^{m_2 - m_1 - 1}(j)\}.$$

This, however, contradicts the definition of $N_{\sigma,j}$ since $m_2 - m_1 < N_{\sigma,j}$.

(ii) Since $\operatorname{card}(O(\sigma, j)) = N_{\sigma,j}$ and by the previous part of the lemma we must have $\sigma^{N_{\sigma,j}}(j) = \sigma^m(j)$ for some $m \in \{0, 1, \dots, N_{\sigma,j} - 1\}$. Thus $\sigma^{N_{\sigma,j} - m}(j) = j$ and so, by the previous part of the lemma we must have $m = 0$. Thus $\sigma^{N_{\sigma,j}}(j) = j$. Now, if $m \in \{1, \dots, N_{\sigma,j} - 1\}$, then

$$\sigma^{N_{\sigma,j}} \circ \sigma^m(j) = \sigma^m \circ \sigma^{N_{\sigma,j}}(j) = \sigma^m(j),$$

giving this part of the lemma.

(iii) Clearly

$$\{j, \sigma(j), \dots, \sigma^{N_{\sigma,j} - 1}(j)\} \subseteq O(\sigma, j).$$

By definition of $N_{\sigma,j}$ and by part (i) equality follows.

(iv) Let $m' \in \{1, \dots, N_{\sigma,j} - 1\}$ and let $j' = \sigma^{m'}(j)$.

$$O(\sigma, j') = \{\sigma^m(j') \mid m \in \mathbb{Z}_{\geq 0}\} = \{\sigma^{m+m'}(j) \mid m \in \mathbb{Z}_{\geq 0}\} \subseteq O(\sigma, j).$$

On the other hand, if $m \in \mathbb{Z}_{>0}$ we can write $m - m' = qN_{\sigma,j} + r$ for $r \in \{0, 1, \dots, N_{\sigma,j} - 1\}$ using the division algorithm. Then

$$\sigma^m(j) = \sigma^{m-m'} \circ \sigma^{m'}(j) = \sigma^r \circ \sigma^{m'}(j) = \sigma^r(j'),$$

and so $O(\sigma, j) \subseteq O(\sigma, j')$. ▼

From the lemma and since the set $\{1, \dots, k\}$ is finite it follows that there exist $j_1, \dots, j_r \in \{1, \dots, k\}$ such that

1. $\{1, \dots, k\} = \cup_{l=1}^r O(\sigma, j_l)$ and
2. $O(\sigma, j_l) \cap O(\sigma, j_m) = \emptyset$ for $l \neq m$.

Let $N_l = \operatorname{card}(O(\sigma, j_l))$ for $l \in \{1, \dots, r\}$. For $l \in \{1, \dots, r\}$ define $\sigma_l \in \mathfrak{S}_k$ by

$$\sigma_l(j) = \begin{cases} \sigma(j), & j \in O(\sigma, j_l), \\ j, & \text{otherwise.} \end{cases}$$

By the lemma we have $\sigma_l = (j_l\ \sigma(j_l)\ \cdots\ \sigma^{N_l - 1}(j_l))$. Moreover, for distinct $l, m \in \{1, \dots, r\}$ the permutations σ_l and σ_m are clearly disjoint. Therefore, by Exercise 4.1.12, the permutations $\sigma_1, \dots, \sigma_l$ commute with one another. We claim that $\sigma = \sigma_1 \circ \cdots \circ \sigma_r$. Indeed, let $j \in \{1, \dots, k\}$ and let $l_j \in \{1, \dots, r\}$ satisfy $j \in O(\sigma, l_j)$. Then, by construction, $\sigma_l(j) = j$ for $l \neq l_j$. We thus have

$$\sigma_1 \circ \cdots \circ \sigma_{l_j} \circ \cdots \circ \sigma_r(j) = \sigma_{l_j} \circ \sigma_1 \circ \cdots \circ \sigma_{l_j - 1} \circ \sigma_{l_j + 1} \circ \cdots \circ \sigma_r(j) = \sigma_{l_j}(j) = \sigma(j),$$

giving the theorem. ■

It is not clear that a permutation cannot be both even and odd, so let us establish this in an illuminating way. In the statement of the result we consider the set $\{-1, 1\}$ to be a group with the product being multiplication in the usual way.

4.1.31 Theorem (The sign homomorphism from the symmetric group) *Let* $k \in \mathbb{Z}_{>0}$*. If* $\sigma \in \mathfrak{S}_k$ *then* σ *is the product of a finite number of transpositions. Moreover, the map* $\mathrm{sign}\colon \mathfrak{S}_k \to \{-1, 1\}$ *given by*

$$\mathrm{sign}(\sigma) = \begin{cases} 1, & \sigma \textit{ is a product of an even number of transpositions}, \\ -1, & \sigma \textit{ is a product of an odd number of transpositions} \end{cases}$$

is a well-defined group homomorphism.

Proof By Theorem 4.1.30 it suffices to show that a cycle is a finite product of adjacent transpositions. However, for a cycle $(j_1 \ \cdots \ j_m)$ we can write

$$(j_1 \ \cdots \ j_m) = (j_1 \ j_2) \cdot (j_1 \ j_3) \cdot \cdots \cdot (j_1 \ j_m),$$

which can be verified directly.

Now we prove that sign is well-defined. Let $\sigma \in \mathfrak{S}_k$. By Theorem 4.1.30 there exist unique (up to order) disjoint cycles $\sigma_1, \dots, \sigma_r$ such that $\sigma = \sigma_1 \circ \cdots \circ \sigma_r$. Let us define $C(\sigma) = r$. In the following lemma we recall the notation $O(\sigma, j)$ introduced in the proof of Theorem 4.1.30.

1 Lemma *Let* $\sigma \in \mathfrak{S}_k$ *and let* $\tau = (j_1, j_2)$*. Then*

(i) $C(\sigma \circ \tau) = C(\sigma) + 1$ *if* $O(\sigma, j_1) = O(\sigma, j_2)$ *and*

(ii) $C(\sigma \circ \tau) = C(\sigma) - 1$ *if* $O(\sigma, j_1) \neq O(\sigma, j_2)$*.*

Proof Suppose that $O(\sigma, j_1) = O(\sigma, j_2)$ and, using the lemma from the proof of Theorem 4.1.30, write

$$O(\sigma, j_1) = \{l_1 = j_1, \dots, l_s = j_2, \dots, l_m\}$$

with $l_p = \sigma^p(l_1)$ for $p \in \{1, \dots, m\}$. Let $\sigma' = (l_1 \ \cdots \ l_p)$. Then we can directly verify that

$$\sigma' \circ \tau = (l_1 \ \cdots \ l_p) \cdot (l_1 \ l_s) = (l_1 \ \cdots \ l_{s-1}) \cdot (l_s \ \cdots \ l_p),$$

giving $\sigma' \circ \tau$ as a product of two cycles. Now note that if j has the property that $O(\sigma, j) \neq O(\sigma, j_1)$ then, using the lemma from the proof of Theorem 4.1.30, $\sigma \circ \tau(j) = \sigma(j)$. Thus $O(\sigma \circ \tau, j) = O(\sigma, j)$ if $j \notin O(\sigma, j_1)$. For $j \in O(\sigma, j_1)$ we have $\sigma(j) = \sigma'(j)$ and also $\sigma \circ \tau(j) = \sigma' \circ \tau(j)$ since $\tau(j) \in O(\sigma, j_1)$. Thus

$$O(\sigma, j_1) = O(\sigma \circ \tau, j_1) \cup O(\sigma \circ \tau, j_2),$$

giving $C(\sigma \circ \tau) = C(\sigma) + 1$.

Now suppose that $O(\sigma, j_1) \neq O(\sigma, j_2)$. Let us write

$$O(\sigma, j_1) = \{j_1, \sigma(j_1), \dots, \sigma^{p_1 - 1}(j_1)\}, \quad O(\sigma, j_2) = \{j_2, \sigma(j_2), \dots, \sigma^{p_2 - 1}(j_2)\}.$$

Let us also define

$$\sigma_1' = (j_1 \ \sigma(j_1) \ \cdots \ \sigma^{p_1 - 1}(j_1)), \quad \sigma_2' = (j_2 \ \sigma(j_2) \ \cdots \ \sigma^{p_2 - 1}(j_2)).$$

One can then directly see that

$$\begin{aligned} \sigma_1' \circ \sigma_2' \circ \tau = (j_1 \ \sigma(j_1) \ \cdots \ \sigma^{p_1 - 1}(j_1)) \cdot (j_2 \ \sigma(j_2) \ \cdots \ \sigma^{p_2 - 1}(j_2)) \cdot (j_1, j_2) \\ = (j_1 \ \sigma(j_1) \ \cdots \ \sigma^{p_1 - 1}(j_1) \ j_2 \ \sigma(j_2) \ \cdots \ \sigma^{p_2 - 1}(j_2)). \end{aligned}$$

Now note that if $j \in O(\sigma, j_1) \cup O(\sigma, j_2)$ then $\sigma(j) = \sigma'_1 \circ \sigma'_2(j)$ whence $\sigma \circ \tau(j) = \sigma'_1 \circ \sigma'_2 \circ \tau(j)$ since $\tau(j) \in O(\sigma, j_1) \cup O(\sigma, j_2)$. Therefore, $O(\sigma, j_1) \cup O(\sigma, j_2) = O(\sigma \circ \tau, j_1)$. Moreover, if $j \notin O(\sigma, j_1) \cup O(\sigma, j_2)$ then obviously $\sigma(j) = \sigma \circ \tau(j)$. Therefore, $O(\sigma \circ \tau, j) = O(\sigma, j)$ in this case. Summarising, $C(\sigma \circ \tau) = C(\sigma) - 1$. ▼

Let $\pi_2 \colon \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ be the canonical projection. Since $\pi_2(m+1) = \pi_2(m-1)$, the lemma shows that $\pi_2(C(\sigma)) = \pi_2(C(\sigma \circ \tau) + 1)$ for every $\sigma \in \mathfrak{S}_k$ and for every transposition τ.

To complete the proof note that $C(e) = k$ if e denotes the identity element of $\mathfrak{S}_k$. Now write $\sigma \in \mathfrak{S}_k$ as a finite product of transpositions: $\sigma = \tau_1 \circ \cdots \circ \tau_p$. Thus

$$\pi_2(C(\sigma)) = \pi_2(C(\tau_1 \circ \cdots \circ \tau_p)) = \pi_2(C(e) + p) = \pi_2(k + p).$$

Note that $\pi_2(C(\sigma))$ is defined independently of the choice of transpositions $\tau_1, \dots, \tau_p$. Thus, if $\sigma = \tau'_1 \circ \cdots \circ \tau'_{p'}$ for transpositions $\tau'_1, \dots, \tau'_{p'}$ then we must have $\pi_2(k+p) = \pi_2(k+p')$ meaning that $\pi_2(p) = \pi_2(p')$. But this means exactly that p and p' are either both even or both odd.

That sign is a homomorphism is a consequence of the obvious fact that the product of even permutations is even, the product of two odd permutations is even, and the product of an even and an odd permutation is odd. ■

Let us give some additional properties of the symmetric group that will be useful to us in our discussions of multilinear maps in Section **??**, derivatives of such maps in Section **??** and Theorem **??**.

Let $k_1, \dots, k_m \in \mathbb{Z}_{\geq 0}$ be such that $\sum_{j=1}^m k_m = k$. Let $\mathfrak{S}_{k_1|\cdots|k_m}$ be the subgroup of $\mathfrak{S}_k$ with the property that elements σ of $\mathfrak{S}_{k_1|\cdots|k_m}$ take the form

$$\begin{pmatrix} 1 & \cdots & k_1 & \cdots & k_1 + \cdots + k_{m-1} + 1 & \cdots & k_1 + \cdots + k_m \\ \sigma_1(1) & \cdots & \sigma_1(k_1) & \cdots & k_1 + \cdots + k_{m-1} + \sigma_m(1) & \cdots & k_1 + \cdots + k_{m-1} + \sigma_m(k_m) \end{pmatrix},$$

where $\sigma_j \in \mathfrak{S}_{k_j}$, $j \in \{1, \dots, m\}$. The assignment $(\sigma_1, \dots, \sigma_m) \mapsto \sigma$ with σ as above is an isomorphism of $\mathfrak{S}_{k_1} \times \cdots \times \mathfrak{S}_{k_m}$***missing stuff*** with $\mathfrak{S}_{k_1|\cdots|k_m}$. Also denote by $\mathfrak{S}_{k_1,\dots,k_m}$ the subset of $\mathfrak{S}_k$ having the property that $\sigma \in \mathfrak{S}_{k_1,\dots,k_m}$ satisfies

$$\sigma(k_1 + \cdots + k_j + 1) < \cdots < \sigma(k_1 + \cdots + k_j + k_{j+1}), \qquad j \in \{0, 1, \dots, m-1\}.$$

Now we have the following result.

4.1.32 Proposition (Decompositions of the symmetric group) *With the above notation, the map $(\sigma_1, \cdots \sigma_{\mathrm{m}}) \mapsto \sigma_1 \circ \cdots \circ \sigma_{\mathrm{m}}$ from $\mathfrak{S}_{\mathrm{k}_1,\dots,\mathrm{k}_\mathrm{m}} \times \mathfrak{S}_{\mathrm{k}_1|\cdots|\mathrm{k}_\mathrm{m}}$ to $\mathfrak{S}_\mathrm{k}$ is a bijection.*

Proof Let P be the set of partitions $(S_1, \dots, S_m)$ of $\{1, \dots, k\}$ (i.e., $\{1, \dots, k\} = \overset{\circ}{\cup}{}_{j=1}^m S_j$) such that $\mathrm{card}(S_j) = k_j$, $j \in \{1, \dots, m\}$. Note that $\mathfrak{S}_k$ acts in a natural way on P. That is, if $(S_1, \dots, S_m) \in P$ and if $\sigma \in \mathfrak{S}_k$ then we can define $\sigma(S_1, \dots, S_m)$ to be the partition $(S'_1, \dots, S'_m) \in P$ for which $\sigma(S_j) = S'_j$ for each $j \in \{1, \dots, m\}$. Now specifically choose $S = (S_1, \dots, S_m) \in P$ by

$$S_j = \{k_0 + \cdots + k_{j-1} + 1, \dots, k_1 + \cdots + k_j\}, \qquad j \in \{1, \dots, m\},$$

taking $k_0 = 0$. Note that $\sigma \in \mathfrak{S}_k$ has the property that $\sigma(S) = S$ if and only if $\sigma \in \mathfrak{S}_{k_1|\cdots|k_m}$. For a general $T = (T_1, \dots, T_m) \in P$ let $\mathfrak{S}_{S \to T}$ be the set of $\sigma \in \mathfrak{S}_k$ that map S to T. Note that for a given $T \in P$ there exists a unique element of $\mathfrak{S}_{k_1,\dots,k_m} \cap \mathfrak{S}_{S\to T}$ (why?). Let us denote this unique permutation by $\sigma_T \in \mathfrak{S}_{k_1,\dots,k_m} \cap \mathfrak{S}_{S\to T}$. We claim that

$$\mathfrak{S}_{S\to T} = \{\sigma_T \circ \sigma' \mid \sigma' \in \mathfrak{S}_{k_1|\cdots|k_m}\}.$$

Indeed, if $\sigma \in \mathfrak{S}_{S\to T}$ then $\sigma_T^{-1} \circ \sigma(S) = S$ and so $\sigma_T^{-1} \circ \sigma = \sigma'$ for some $\sigma' \mathfrak{S}_{k_1|\cdots|k_m}$. Thus $\sigma = \sigma_T \circ \sigma'$ and so

$$\mathfrak{S}_{S\to T} \subseteq \{\sigma_T \circ \sigma' \mid \sigma' \in \mathfrak{S}_{k_1|\cdots|k_m}\}.$$

Conversely, if $\sigma' \in \mathfrak{S}_{k_1|\cdots|k_m}$ then $\sigma_T \circ \sigma' \in \mathfrak{S}_{S\to T}$ since $\sigma'(S) = S$. This gives $\mathfrak{S}_{S\to T} = \sigma_T \mathfrak{S}_{k_1|\cdots|k_m}$. Since σ_T is the unique element of $\mathfrak{S}_{k_1,\dots,k_m}$ for which this holds, it follows that if $\sigma \in \mathfrak{S}_{S\to T}$ for some $T \in P$ we have $\sigma = \sigma_1 \circ \sigma_2$ for unique $\sigma_1 \in \mathfrak{S}_{k_1,\dots,k_m}$ and $\sigma_2 \in \mathfrak{S}_{k_1|\cdots|k_m}$. Now, if $\sigma \in \mathfrak{S}_k$ then $\sigma \in \mathfrak{S}_{S\to T}$ for $T = \sigma^{-1}(S)$, and so the result holds. ■

Exercises

4.1.1 Do the following;

(a) prove Proposition 4.1.6;

(b) state which of the statements in Proposition 4.1.6 holds for semigroups;

(c) state which of the statements in Proposition 4.1.6 holds for monoids.

4.1.2 Let M be a monoid for which $ab = ba$ for all $a, b \in \mathsf{M}$, and let $m_1, \dots, m_k \in \mathsf{M}$ be elements for which there exists no inverse. Show that there is also no inverse for $m_1 \cdots m_k$.

4.1.3 Let G be a group and let $a, b, c \in \mathsf{G}$.

(a) Show that if $ab = ac$ then $b = c$.

(b) Show that if $ac = bc$ then $a = b$.

4.1.4 Let G and H be groups. Show that, if $\phi\colon \mathsf{G} \to \mathsf{H}$ is an isomorphism, then ϕ^{-1} is a homomorphism, and so also an isomorphism.

4.1.5 Prove Proposition 4.1.10.

4.1.6 Show that the following sets are subgroups of $\mathbb{R}$ with the group operation of addition:

(a) $\mathbb{Z}$;

(b) $\mathbb{Z}(\Delta) = \{j\delta \mid j \in \mathbb{Z}\}$;

(c) $\mathbb{Q}$.

The next two parts of this problem suppose that you know something about polynomials; we consider these in detail in Section **??**. In any case, you should also show that the following sets are subgroups of $\mathbb{R}$ with the group operation of addition.

(d) the set $\bar{\mathbb{Q}} \cap \mathbb{R}$ of real algebraic numbers (recall that $z \in \mathbb{C}$ is an ***algebraic number*** if there exists a polynomial $P \in \mathbb{Z}[\xi]$ (i.e., one with integer coefficients) for which $P(z) = 0$, and we denote the set of algebraic numbers by $\bar{\mathbb{Q}}$);

(e) the set $\mathbb{K} \cap \mathbb{R}$ of real algebraic integers (recall that $z \in \mathbb{C}$ is an ***algebraic integer*** if there exists a monic polynomial $P \in \mathbb{Z}[\xi]$ (i.e., one with integer coefficients, and with the highest degree coefficient being 1) for which $P(z) = 0$, and we denote the set of algebraic integers by $\bar{\mathbb{K}}$).

4.1.7 Show that the subsets

(a) $x_0 + \mathbb{Q} = \{x_0 + q \mid q \in \mathbb{Q}\}$ for $x_0 \in \mathbb{R}$ and

(b) $\mathbb{Z}(x_0, \Delta) = \{x_0 + k\Delta \mid k \in \mathbb{Z}\}$ for $x_0 \in \mathbb{R}$

of $\mathbb{R}$ are semigroups with the binary operation

$$(x_0 + y_1) + (x_0 + y_2) = x_0 + y_1 + y_2.$$

Answer the following questions.

(c) Show that $x_0 + \mathbb{Q} = \mathbb{Q}$ if and only if $x_0 \in \mathbb{Q}$ and that $\mathbb{Z}(x_0, \Delta) = \mathbb{Z}(\Delta)$ if and only if $x_0 \in \mathbb{Z}(\Delta)$.

(d) Suppose that the binary operations on the semigroups $x_0+\mathbb{Q}$ and $\mathbb{Z}(x_0, \Delta)$ are as defined above. Show that the semigroup is a subgroup of $\mathbb{R}$ if and only if $x_0 = 0$.

4.1.8 Show that N is a normal subgroup of G if and only if $gng^{-1} \in \mathsf{N}$ for all $g \in \mathsf{G}$ and $n \in \mathsf{N}$.

4.1.9 Let G and H be groups and let $\phi\colon \mathsf{G} \to \mathsf{H}$ be an epimorphism. Show that the map $\phi_0\colon \mathsf{G}/\ker(\phi) \to \mathsf{H}$ defined by $\phi_0(g \ker(\phi)) = \phi(g)$ is a well-defined isomorphism.

In the following exercise you will use the definition that a transposition $\sigma \in \mathfrak{S}_k$ is ***adjacent*** if it has the form $\sigma = (j, j+1)$ for some $j \in \{1, \dots, k-1\}$. •

4.1.10 Show that any permutation $\sigma \in \mathfrak{S}_k$ is a finite product of adjacent transpositions.

4.1.11 Show that the only permutation that is a cycle of length 1 is the identity map.

4.1.12 Show that if $\sigma_1, \sigma_2 \in \mathfrak{S}_k$ are disjoint then $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$.

Section 4.2

Fields

In this section we consider a special sort of ring, one whose nonzero elements are units. These special rings, called fields, are important to us because they form the backdrop for linear algebra, and as such are distinguished in the set of rings.

Do I need to read this section? Readers who are familiar with the basic arithmetic properties of real and numbers can probably omit reading this section. Certain of the ideas we discuss here will be important in our discussion of polynomials in Section **??**, and so a reader wishing to learn about polynomials might benefit from first understanding fields in the degree of generality we present them in this section. •

4.2.1 Definitions and basic properties

The definition of a field proceeds easily once one has on hand the notion of a ring. However, in our definition we repeat the basic axiomatic structure so a reader will not have to refer back to Definition **??**.

4.2.1 Definition A ***division ring*** is a unit ring in which every nonzero element is a unit, and a ***field*** is a commutative division ring. Thus a field is a set F with two binary operations, $(a_1, a_2) \mapsto a_1 + a_2$ and $(a_1, a_2) \mapsto a_1 \cdot a_2$, called ***addition*** and ***multiplication***, respectively, and which together satisfy the following rules:

(i) $(a_1 + a_2) + a_3 = a_1 + (a_2 + a_3)$, $a_1, a_2, a_3 \in \mathsf{F}$ (***associativity*** of addition);

(ii) $a_1 + a_2 = a_2 + a_1$, $a_1, a_2 \in \mathsf{F}$ (***commutativity*** of addition);

(iii) there exists $0_{\mathsf{F}} \in \mathsf{F}$ such that $a + 0_{\mathsf{F}} = a$, $a \in \mathsf{F}$ (***additive identity***);

(iv) for $a \in \mathsf{F}$, there exists $-a \in \mathsf{F}$ such that $a + (-a) = 0_{\mathsf{F}}$ (***additive inverse***);

(v) $(a_1 \cdot a_2) \cdot a_3 = a_1 \cdot (a_2 \cdot a_3)$, $a_1, a_2, a_3 \in \mathsf{F}$ (***associativity*** of multiplication);

(vi) $a_1 \cdot a_2 = a_2 \cdot a_1$, $a_1, a_2 \in \mathsf{F}$ (***commutativity*** of multiplication);

(vii) $a_1 \cdot (a_2 + a_3) = (a_1 \cdot a_2) + (a_1 \cdot a_3)$, $a_1, a_2, a_3 \in \mathsf{F}$ (***left distributivity***);

(viii) there exists $1_{\mathsf{F}} \in \mathsf{F}$ such that $1_{\mathsf{F}} \cdot a = a$, $a \in \mathsf{F}$ (***multiplicative identity***);

(ix) for $a \in \mathsf{F}$, there exists $a^{-1} \in \mathsf{F}$ such that $a^{-1} \cdot a = 1_{\mathsf{F}}$ (***multiplicative inverse***);

(x) $(a_1 + a_2) \cdot a_3 = (a_1 \cdot a_3) + (a_2 \cdot a_3)$, $a_1, a_2, a_3 \in \mathsf{F}$ (***right distributivity***). •

The following result gives some properties of fields that follow from the definitions or which follow from general properties of rings.

4.2.2 Proposition (Basic properties of fields) *Let F be a field and denote $\mathsf{F}^* = \mathsf{F} \setminus \{0_{\mathsf{F}}\}$. Then the following statements hold:*

(i) F^*, equipped with the binary operation of multiplication, is a group;

(ii) F is an integral domain;

(iii) F *is a Euclidean domain;*

(iv) F *is a principal ideal domain;*

(v) F *is a unique factorisation domain.*

4.2.3 Remark (Fields as unique factorisation domains) It is worth commenting on the nature of fields as unique factorisation domains. The definition of a unique factorisation domain requires that one be able to factor nonzero nonunits as products of irreducibles. However, in fields there are neither any nonzero nonunits, nor any irreducibles. Therefore, fields are vacuous unique factorisation domains. •

Let us give some examples of fields.

4.2.4 Examples (Fields)

1. $\mathbb{Z}$ is not a field since the only units are -1 and 1.
2. $\mathbb{Q}$ is a field.
3. $\mathbb{R}$ is a field.
4. The ring $\mathbb{Z}_k$ is a field if and only if k is prime. This follows from our discussion in Example **??–??** of the units in $\mathbb{Z}_k$. However, let us repeat the argument here, using Bézout's Identity in a coherent manner. We rely on the fact that $\mathbb{Z}$ is a Euclidean domain (Theorem **??**), and so a principal ideal domain (Theorem **??**), and so a unique factorisation domain (Theorem **??**).

   Suppose that k is prime and let $j \in \{1, \dots, k-1\}$. Then 1 is a greatest common divisor for $\{j, k\}$, and by Corollary **??** this means that there exists $l, m \in \mathbb{Z}$ such that $lj + mk = 1$. Therefore, $(j + k\mathbb{Z})(l + k\mathbb{Z}) = lj + k\mathbb{Z} = 1 + k\mathbb{Z}$, and so $j + k\mathbb{Z}$ is a unit.

   Now suppose that $\mathbb{Z}_k$ is a field and let $j \in \{1, \dots, k-1\}$. Then there exists $l \in \{1, \dots, k-1\}$ such that $(j+k\mathbb{Z})(l+k\mathbb{Z}) = 1+k\mathbb{Z}$. Therefore, $jl+mk = 1$ for some $m \in \mathbb{Z}$, and by Corollary **??** we can conclude that j and k are relatively prime. Since this must hold for every $j \in \{1, \dots, k-1\}$, it follows from Proposition **??** that k is prime. •

4.2.2 Fraction fields

Corresponding to a commutative unit ring is a natural field given by "fractions" in R. The construction here strongly resembles the construction of the rational numbers from the integers, so readers may wish to review Section 2.1.1.

4.2.5 Definition (Fraction field) Let R be an integral domain and define an equivalence relation $\sim$ in $\mathsf{R} \times (\mathsf{R} \setminus \{0_\mathsf{R}\})$ by

$$(r, s) \sim (r', s') \qquad \Longleftrightarrow \qquad rs' - r's = 0_\mathsf{R}$$

(the reader may verify in Exercise 4.2.1 that $\sim$ is indeed an equivalence relation). The set of equivalence classes under this equivalence relation is the ***fraction field*** of R, and is denoted by F_R. The equivalence class of (r, s) is denoted by $\frac{r}{s}$. •

Let us show that the name fraction *field* is justified.

4.2.6 Theorem (The fraction field is a field) *If* R *is an integral domain, then* F_R *is a field when equipped with the binary operations of addition and multiplication defined by*

$$\frac{r_1}{s_2} + \frac{r_2}{s_2} = \frac{r_1 s_2 + r_2 s_1}{s_1 s_1}, \quad \frac{r_1}{s_1} \cdot \frac{r_1 \cdot r_2}{s_1 \cdot s_2}.$$

Moreover, the map $r \mapsto \frac{r}{1_\mathsf{R}}$ *is a ring monomorphism from* R *to* F_R.

Proof If one defines the zero element in the field to be $\frac{0_\mathsf{R}}{1_\mathsf{R}}$, the unity element to be $\frac{1_\mathsf{R}}{1_\mathsf{R}}$, the additive inverse of $\frac{r}{s}$ to be $\frac{-r}{s}$, and the multiplicative inverse of $\frac{r}{s}$ to be $\frac{s}{r}$, then it is a matter of tediously checking the conditions of Definition 4.2.1 to see that F_R is a field. The final assertion is also easily checked. We leave the details of this to the reader as Exercise 4.2.2. ■

The only interesting example of a fraction field that we have encountered thus is the field $\mathbb{Q}$ which is obviously the fraction field of $\mathbb{Z}$. In Section **??** we will encounter the field of rational functions that is associated with a polynomial ring.

4.2.3 Subfields, field homomorphisms, and characteristic

All of the ideas in this section have been discussed in the more general setting of rings in Section **??**. Therefore, we restrict ourselves to making the (obvious) definitions and pointing out the special features arising when one restricts attention to fields.

Since fields are also rings, the following definition is the obvious one.

4.2.7 Definition (Subfield) A nonempty subset K of a field F is a ***subfield*** if K is a subring of the ring F that (1) contains 1_F and (2) contains a^{-1} for every $a \in \mathsf{K} \setminus \{0_\mathsf{F}\}$. •

Of course, just as in Definition **??**, a subset $\mathsf{K} \subseteq \mathsf{F}$ is a subfield if and only if (1) $a_1 + a_2 \in \mathsf{K}$ for all $a_1, a_2 \in \mathsf{K}$, (2) $a_1 \cdot a_2 \in \mathsf{K}$ for all $a_1, a_2 \in \mathsf{K}$, (3) $-a \in \mathsf{K}$ for all $a \in \mathsf{K}$, (4) $1_\mathsf{F} \in \mathsf{K}$, and (4) $a^{-1} \in \mathsf{K}$ for all nonzero $a \in \mathsf{K}$. Note that we do require that 1_F be an element of a subfield so as to ensure that subfields are actually fields (see Exercises 4.2.3 and 4.2.4).

Note that we have not made special mention of ideals which were so important to our characterisations of rings. The reason for this is that ideals for fields are simply not very interesting, as the following result suggests.

4.2.8 Proposition (Ideals of fields) *If* R *is a commutative unit ring with more than one element, then the following statements are equivalent:*

(i) R *is a field;*

(ii) $\{0_\mathsf{R}\}$ *is a maximal ideal of* R*;*

(iii) *if* I *is an ideal of* R, *then either* $\mathsf{I} = \{0_\mathsf{R}\}$ *or* $\mathsf{I} = \mathsf{R}$.

Proof (i) $\Longrightarrow$ (ii) Suppose that I is an ideal of R for which $\{0_\mathsf{R}\} \subseteq \mathsf{I}$. If $\{0_\mathsf{R}\} \neq \mathsf{I}$ then let $a \in \mathsf{I} \setminus \{0_\mathsf{R}\}$. For any $r \in \mathsf{R}$ we then have $r = (ra^{-1})a$, meaning that $r \in \mathsf{I}$. Thus $\mathsf{I} = \mathsf{R}$, and so $\{0_\mathsf{R}\}$ is maximal.

(ii) $\Longrightarrow$ (iii) This follows immediately by the definition of maximal ideal.

(iii) $\Longrightarrow$ (i) Let $r \in \mathsf{R} \setminus \{0_\mathsf{R}\}$ and consider the ideal (r). Since $(r) \neq \{0_\mathsf{R}\}$ we must have $(r) = \mathsf{R}$. In particular, $1_\mathsf{F} = rs$ for some $s \in \mathsf{R}$, and so r is a unit. ■

The interesting relationship between fields and ideals, then, does not come from considering ideals of fields. However, there is an interesting connection of fields to ideals. This connection, besides being of interest to us in Section **??**, gives some additional insight to the notion of maximal ideals. The result mirrors that for prime ideals given as Theorem **??**.

4.2.9 Theorem (Quotients by maximal ideals are fields, and vice versa) *If* R *is a commutative unit ring with more than one element and if* $\mathsf{I} \subseteq \mathsf{R}$ *is an ideal, then the following two statements are equivalent:*

(i) I *is a maximal ideal;*

(ii) R/I *is a field.*

Proof Denote by $\pi_{\mathsf{I}} \colon \mathsf{R} \to \mathsf{R}/\mathsf{I}$ the canonical projection. Suppose that I is a maximal ideal and let $\mathsf{J} \subseteq \mathsf{R}/\mathsf{I}$ be an ideal. We claim that

$$\tilde{\mathsf{J}} = \{r \in \mathsf{R} \mid \pi_{\mathsf{I}}(r) \in \mathsf{J}\}$$

is an ideal in R. Indeed, let $r_1, r_2 \in \tilde{\mathsf{J}}$ and note that $\pi_{\mathsf{I}}(r_1 - r_2) = \pi_{\mathsf{I}}(r_1) - \pi_{\mathsf{I}}(r_2) \in \mathsf{J}$ since π_{I} is a ring homomorphism and since J is an ideal. Thus $r_1 - r_2 \in \tilde{\mathsf{J}}$. Now let $r \in \tilde{\mathsf{J}}$ and $s \in \mathsf{R}$ and note that $\pi_{\mathsf{I}}(sr) = \pi_{\mathsf{I}}(s)\pi_{\mathsf{I}}(r) \in \mathsf{J}$, again since π_{I} is a ring homomorphism and since J is an ideal. Thus $\tilde{\mathsf{J}}$ is an ideal. Clearly $\mathsf{I} \subseteq \tilde{\mathsf{J}}$ so that either $\tilde{\mathsf{J}} = \mathsf{I}$ or $\tilde{\mathsf{J}} = \mathsf{R}$. In the first case $\mathsf{J} = \{0_{\mathsf{R}} + \mathsf{I}\}$ and in the second case $\mathsf{J} = \mathsf{R}/\mathsf{I}$. Thus the only ideals of R/I are $\{0_{\mathsf{R}} + \mathsf{I}\}$ and R/I. That R/I is a field follows from Proposition 4.2.8.

Now suppose that R/I is a field and let J be an ideal of R for which $\mathsf{I} \subseteq \mathsf{J}$. We claim that $\pi_{\mathsf{I}}(\mathsf{J})$ is an ideal of R/I. Indeed, let $r_1 + \mathsf{I}, r_2 + \mathsf{I} \in \pi_{\mathsf{I}}(\mathsf{J})$. Then $r_1, r_2 \in \mathsf{J}$ and so $r_1 - r_2 \in \mathsf{J}$, giving $(r_1 - r_2) + \mathsf{I} \in \pi_{\mathsf{I}}(\mathsf{J})$. If $r + \mathsf{I} \in \pi_{\mathsf{I}}(\mathsf{J})$ and if $s + \mathsf{I} \in \mathsf{R}/\mathsf{I}$, then $r \in \mathsf{J}$ and so $sr \in \mathsf{J}$. Then $sr + \mathsf{I} \in \pi_{\mathsf{I}}(\mathsf{J})$, thus showing that $\pi_{\mathsf{I}}(\mathsf{J})$ is indeed an ideal. Since R/I is a field, by Proposition 4.2.8 we may conclude that either $\pi_{\mathsf{I}}(\mathsf{J}) = \{0_{\mathsf{R}} + \mathsf{I}\}$ or that $\pi_{\mathsf{I}}(\mathsf{J}) = \mathsf{R}/\mathsf{I}$. In the first case we have $\mathsf{J} \subseteq \mathsf{I}$ and hence $\mathsf{J} = \mathsf{I}$, and in the second case we have $\mathsf{J} = \mathsf{R}$. Thus I is maximal. ■

The definition of a homomorphism of fields follows from the corresponding definition for rings.

4.2.10 Definition (Field homomorphism, epimorphism, monomorphism, and isomorphism) For fields F and K, a map $\phi \colon \mathsf{F} \to \mathsf{K}$ is a ***field homomorphism*** (resp. ***epimorphism, monomorphism, isomorphism***) if it is a homomorphism (resp. epimorphism, monomorphism, isomorphism) of rings. If there exists an isomorphism from F to K, then F and K are ***isomorphic***. •

The definitions of kernel and image for field homomorphisms are then special cases of the corresponding definitions for rings, and the corresponding properties also follow, just as for rings.

For fields one adopts the notion of characteristic from rings. Thus a field has ***characteristic* k** if it has characteristic k as a ring. The next result gives the analogue of Proposition **??** for fields.

4.2.11 Proposition (Property of fields with given characteristic) *If* F *is a field then the following statements hold:*

(i) *if* F *has characteristic zero then there exists a subfield* K *of* F *that is isomorphic to* $\mathbb{Q}$*;*

(ii) *if* F *has characteristic* $\mathrm{k} \in \mathbb{Z}_{>0}$ *then* k *is prime and there exists a subfield* K *of* F *that is isomorphic to* $\mathbb{Z}_{\mathrm{k}}$*.*

Proof First suppose that F has characteristic zero. As in the proof of Proposition **??**, let $\phi\colon \mathbb{Z} \to \mathsf{R}$ be the map $\phi(j) = j1_{\mathsf{F}}$, and recall that this map is a monomorphism, and so an isomorphism from $\mathbb{Z}$ to image(ϕ). For $j1_{\mathsf{F}} \in \mathrm{image}(\phi) \setminus \{0_{\mathsf{F}}\}$, since F is a field there exists $(j1_{\mathsf{F}})^{-1} \in \mathsf{F}$ such that $(j1_{\mathsf{F}}) \cdot (j1_{\mathsf{F}})^{-1} = 1_{\mathsf{F}}$. We map then define a map $\bar{\phi}\colon \mathbb{Q} \to \mathsf{F}$ by $\bar{\phi}(\frac{j}{k}) = (j1_{\mathsf{F}})(k1_{\mathsf{F}})^{-1}$. First let us show that this map is well defined. Suppose that $\frac{j_1}{k_1} = \frac{j_2}{k_2}$, or equivalently that $j_1k_2 = j_2k_1$. Then, using Proposition **??**,

$$(j_1 1_{\mathsf{F}})(k_2 1_{\mathsf{F}}) = 1_{\mathsf{F}}(j_1(k_2 1_{\mathsf{F}})) = (j_1k_2)1_{\mathsf{F}} = (j_2k_1)1_{\mathsf{F}} = 1_{\mathsf{F}}(j_2(k_1 1_{\mathsf{F}})) = (j_2 1_{\mathsf{F}})(k_1 1_{\mathsf{F}}).$$

Thus $(j_1 1_{\mathsf{F}})(k_1 1_{\mathsf{F}})^{-1} = (j_2 1_{\mathsf{F}})(k_2 1_{\mathsf{F}})^{-1}$, and so $\bar{\phi}(\frac{j_1}{k_1}) = \bar{\phi}(\frac{j_2}{k_2})$. Now let us show that $\bar{\phi}$ is a monomorphism. Suppose that $(j_1 1_{\mathsf{F}})(k_1 1_{\mathsf{F}})^{-1} = (j_2 1_{\mathsf{F}})(k_2 1_{\mathsf{F}})^{-1}$ so that, using Proposition **??**, $(j_1k_2 - j_2k_1)1_{\mathsf{F}} = 0_{\mathsf{F}}$. Then it follows that $\frac{j_1}{k_1} = \frac{j_2}{k_2}$ since F has characteristic zero. Next we show that $\bar{\phi}$ is a homomorphism. We compute, after an application of Proposition **??**,

$$\bar{\phi}(\tfrac{j_1}{k_1} + \tfrac{j_2}{k_2}) = ((j_1k_2 + j_2k_1)1_{\mathsf{F}})(k_1k_2 1_{\mathsf{F}})^{-1} = (j_1k_2 1_{\mathsf{F}})(k_1k_2 1_{\mathsf{F}})^{-1} + (j_2k_1 1_{\mathsf{F}})(k_1k_2 1_{\mathsf{F}})^{-1}.$$

Another application of Proposition **??** gives

$$(k_1 1_{\mathsf{F}})(k_2 1_{\mathsf{F}})\bar{\phi}(\tfrac{j_1}{k_1} + \tfrac{j_2}{k_2}) = (j_1k_2 1_{\mathsf{F}}) + (j_2k_1 1_{\mathsf{F}}),$$

which in turn gives

$$\bar{\phi}(\tfrac{j_1}{k_1} + \tfrac{j_2}{k_2}) = (k_1 1_{\mathsf{F}})^{-1}(k_2 1_{\mathsf{F}})((j_1k_2 1_{\mathsf{F}}) + (j_2k_1 1_{\mathsf{F}})) = (j_1 1_{\mathsf{F}})(k_1 1_{\mathsf{F}})^{-1} + (j_2 1_{\mathsf{F}})(k_2 1_{\mathsf{F}})^{-1},$$

or $\bar{\phi}(\frac{j_1}{k_1} + \frac{j_2}{k_2}) = \bar{\phi}(\frac{j_1}{k_1}) + \bar{\phi}(\frac{j_2}{k_2})$. We also have

$$\bar{\phi}(\tfrac{j_1}{k_1}\tfrac{j_2}{k_2}) = (j_1j_2 1_{\mathsf{F}})(k_1k_2 1_{\mathsf{F}})^{-1},$$

which gives, in turn,

$$(k_1 1_{\mathsf{F}})(k_2 1_{\mathsf{F}})\bar{\phi}(\tfrac{j_1}{k_1}\tfrac{j_2}{k_2}) = (j_1 1_{\mathsf{F}})(j_2 1_{\mathsf{F}})$$

and

$$\bar{\phi}(\tfrac{j_1}{k_1}\tfrac{j_2}{k_2}) = (j_1 1_{\mathsf{F}})(k_1 1_{\mathsf{F}})^{-1}(j_2 1_{\mathsf{F}})(k_2 1_{\mathsf{F}})^{-1},$$

or $\bar{\phi}(\frac{j_1}{k_1}\frac{j_2}{k_2}) = \bar{\phi}(\frac{j_1}{k_1})\bar{\phi}(\frac{j_2}{k_2})$. Thus image($\bar{\phi}$) is a subfield of F isomorphic to $\mathbb{Q}$ by the isomorphism $\bar{\phi}$.

For the second part of the result, suppose that $k = k_1k_2$ for $k_1, k_2 \in \{2, \dots, k-1\}$. Then, if F has characteristic k we have

$$0 = k1_{\mathsf{F}} = (k_1k_2)1_{\mathsf{F}} = (k_1 1_{\mathsf{F}})(k_2 1_{\mathsf{F}}).$$

Since F is an integral domain this means, by Exercise **??**, that either $k_1 1_{\mathsf{F}} = 0$ or $k_2 1_{\mathsf{F}} = 0$. This contradicts the fact that F has characteristic k, and so it must not be possible to factor k as a product of positive integers in $\{2, \dots, k-1\}$. Thus k is prime. That F contains a subfield that is isomorphic to $\mathbb{Z}_k$ follows from Proposition **??**. ■

We note that the construction in the proof of a subfield K isomorphic to $\mathbb{Q}$ or $\mathbb{Z}_k$ is explicit, and is by construction the smallest subfield of F. This subfield has a name.

4.2.12 Definition (Prime field) For a field F, the smallest subfield of F is the ***prime field*** of F and is denoted by F_0. •

Exercises

4.2.1 Show that the relation $\sim$ of Definition 4.2.5 is an equivalence relation.

4.2.2 Prove Theorem 4.2.6.

4.2.3 Give a subring of $\mathbb{R}$ that is not a subfield.

4.2.4 Show that, if K is a subfield of F, then K is a field using the binary operations of addition and multiplication of F, restricted to K.

4.2.5 Let F be a field with $\mathsf{K} \subseteq \mathsf{F}$. Show that K is a subfield if and only if

1. $1_{\mathsf{F}} \in \mathsf{K}$,
2. $a - b \in \mathsf{K}$ for each $a, b \in \mathsf{K}$, and
3. $ab^{-1} \in \mathsf{K}$ for each $a, b \in \mathsf{K}$ with $b \neq 0_{\mathsf{F}}$.

Section 4.3

Vector spaces

One of the more important structures that we will use at a fairly high degree of generality is that of a vector space. As with almost everything we have encountered in this chapter, a vector space is a set equipped with certain operations. In the case of vector spaces, one of these operations melds the vector space together with another algebraic structure, in this case a field. A typical first encounter with vector spaces deals primarily with the so-called finite-dimensional case. In this case, a great deal, indeed, pretty much everything, can be said about the structure of these vector spaces. However, in these volumes we shall also encounter so-called infinite-dimensional vector spaces. A study of the structure of these gets rather more detailed than the finite-dimensional case. In this section we deal only with algebraic matters. Important additional structure in the form of a topology is the topic of Chapter **??**.

Do I need to read this section? If you are not already familiar with the idea of an abstract vector space, then you need to read this section. If you are, then it can be bypassed, and perhaps referred to as needed. Parts of this section are also good ones for readers looking for simple proofs that illustrate certain techniques for proving things. These ceases to become true when we discuss bases, since we take an abstract approach motivated by the fact that many of the vector spaces we deal with in these volumes are infinite-dimensional. •

4.3.1 Definitions and basic properties

Throughout this section we let F be a general field, unless otherwise stated. The fields of most interest to us will be $\mathbb{R}$ (see Section 2.1) and $\mathbb{C}$ (see Section **??**). However, most constructions done with vector spaces are done just as conveniently for general fields as for specific ones.

4.3.1 Definition (Vector space) Let F be a field. A ***vector space*** over F, or an **F-*vector space***, is a nonempty set V with two operations: (1) ***vector addition***, denoted by $\mathsf{V} \times \mathsf{V} \ni (v_1, v_2) \mapsto v_1 + v_2 \in \mathsf{V}$, and (2) ***scalar multiplication***, denoted by $\mathsf{F} \times \mathsf{V}(a, v) \mapsto av \in \mathsf{V}$. Vector addition and scalar multiplication must satisfy the following rules:

(i) $v_1 + v_2 = v_2 + v_1$, $v_1, v_2 \in \mathsf{V}$ (***commutativity***);

(ii) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$, $v_1, v_2, v_3 \in \mathsf{V}$ (***associativity***);

(iii) there exists an vector $0_{\mathsf{V}} \in \mathsf{V}$ with the property that $v + 0_{\mathsf{V}} = v$ for every $v \in \mathsf{V}$ (***zero vector***);

(iv) for every $v \in \mathsf{V}$ there exists a vector $-v \in \mathsf{V}$ such that $v + (-v) = 0_{\mathsf{V}}$ (***negative vector***);

(v) $a(bv) = (ab)v$, $a, b \in \mathsf{F}$, $v \in \mathsf{V}$ (***associativity***);

(vi) $1_{\mathsf{F}}v = v$, $v \in \mathsf{V}$;

(vii) $a(v_1 + v_2) = av_1 + av_2$, $a \in \mathsf{F}$, $v_1, v_2 \in \mathsf{V}$ (***distributivity***);

(viii) $(a_1 + a_2)v = a_1v + a_2v$, $a_1, a_2 \in \mathsf{F}$, $v \in \mathsf{V}$ (***distributivity*** again).

A ***vector*** in a vector space V is an element of V. •

We have already encountered some examples of vector spaces. Let us indicate what some of these are, as well as introduce some important new examples of vector spaces. The verifications that the stated sets are vector spaces is routine, and we leave this to the reader in the exercises.

4.3.2 Examples (Vector spaces)

1. Consider a set $0_{\mathsf{V}} = \{v\}$ with one element. There are no choices for the F-vector space structure in this case. We must have $v + v = v$, $av = v$ for every $a \in \mathsf{F}$, $-v = v$, and $0_{\mathsf{V}} = v$. One can then verify that $\{v\}$ is then indeed an F-vector space. This vector space is called the ***trivial vector space***, and is sometimes denoted by $\{0\}$, reflecting the fact that the only vector in the vector space is the zero vector.
2. Let F^n denote the n-fold Cartesian product of F with itself. Let us denote a typical element of F^n by $(v_1, \dots, v_n)$. We define vector addition in F^n by
$$(u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n)$$
and we define scalar multiplication in F^n by
$$a(v_1, \dots, v_n) = (av_1, \dots, av_n).$$
The vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, over $\mathbb{R}$ and $\mathbb{C}$, respectively, will be of particular importance to us. The reader who has no previous knowledge of vector spaces would be well served by spending some time understanding the geometry of vector addition and scalar multiplication in, say, $\mathbb{R}^2$.
3. Let us denote by F^∞ the set of sequences in F. Thus an element of F^∞ is a sequence $(a_j)_{j\in\mathbb{Z}_{>0}}$ with $a_j \in \mathsf{F}$, $j \in \mathbb{Z}_{>0}$. We define vector addition and scalar multiplication by
$$(a_j)_{j\in\mathbb{Z}_{>0}} + (b_j)_{j\in\mathbb{Z}_{>0}} = (a_j + b_j)_{j\in\mathbb{Z}_{>0}}, \quad a(a_j)_{j\in\mathbb{Z}_{>0}} = (aa_j)_{j\in\mathbb{Z}_{>0}},$$
respectively. This can be verified to make F^∞ into an F-vector space. It is tempting to think of things like $\mathsf{F}^\infty = \lim_{n\to\infty} \mathsf{F}^n$, but one must exercise care, since the limit needs definition. This is the realm of Chapter **??**.
4. Let us denote by F_0^∞ the subset of F^∞ consisting of sequences for which all but a finite number of terms is zero. Vector addition and scalar multiplication are defined for F_0^∞ are defined just as for F^∞. It is just as straightforward to verify that these operations make F_0^∞ an F-vector space.
5. If K is a field extension of F (see Definition **??**) and if V is a K-vector space, then V is also an F-vector space with the operation of vector addition being exactly that of V as a K-vector space, and with scalar multiplication simply being the restriction of scalar multiplication by K to F.

6. The set $\mathsf{F}[\xi]$ of polynomials over F is an F-vector space. Vector addition is addition in the usual sense of polynomials, and scalar multiplication is multiplication of polynomials, using the fact that F is a subring of $\mathsf{F}[\xi]$ consisting of the constant polynomials.
7. Denote by $\mathsf{F}_k[\xi]$ the polynomials over F of degree at most k. Using the same definitions of vector addition and scalar multiplication as were used for the F-vector space $\mathsf{F}[\xi]$ in the preceding example, $\mathsf{F}_k[\xi]$ is an F-vector space.
8. Let S be a set and, as in Definition 1.3.1, let F^S be the set of maps from S to F. Let us define vector addition and scalar multiplication in F^S by

$$(f+g)(x) = f(x) + g(x), \quad (af)(x) = a(f(x))$$

for $f, g \in \mathsf{F}^S$ and $a \in \mathsf{F}$. One may directly verify that these operations indeed satisfy the conditions to make F^S into an F-vector space.
9. Let $I \subseteq \mathbb{R}$ be an interval and let $\mathsf{C}^0(I;\mathbb{R})$ denote the set of continuous $\mathbb{R}$-valued functions on I. Following the preceding example, define vector addition and scalar multiplication in $\mathsf{C}^0(I;\mathbb{R})$ by

$$(f+g)(x) = f(x) + g(x), \quad (af)(x) = a(f(x)), \qquad f, g \in \mathsf{C}^0(I;\mathbb{R}),\ a \in \mathbb{R},$$

respectively. With these operations, one can verify that $\mathsf{C}^0(I;\mathbb{R})$ is a $\mathbb{R}$-vector space. •

Let us now prove some elementary facts about vector spaces.

4.3.3 Proposition (Properties of vector spaces) *Let F be a field and let V be an F-vector space. The following statements hold:*

(i) there exists exactly one vector $0_{\mathsf{V}} \in \mathsf{V}$ such that $\mathrm{v} + 0_{\mathsf{V}} = \mathrm{v}$ for all $\mathrm{v} \in \mathsf{V}$;

(ii) for each $\mathrm{v} \in \mathsf{V}$ there exists exactly one vector $-\mathrm{v} \in \mathsf{V}$ such that $\mathrm{v} + (-\mathrm{v}) = 0_{\mathsf{V}}$;

(iii) $\mathrm{a}0_{\mathsf{V}} = 0_{\mathsf{V}}$ for all $\mathrm{a} \in \mathsf{F}$;

(iv) $0_{\mathsf{F}}\mathrm{v} = 0_{\mathsf{V}}$ for each $\mathrm{v} \in \mathsf{V}$;

(v) $\mathrm{a}(-\mathrm{v}) = (-\mathrm{a})\mathrm{v} = -(\mathrm{av})$ for all $\mathrm{a} \in \mathsf{F}$ and $\mathrm{v} \in \mathsf{V}$;

(vi) if $\mathrm{av} = 0_{\mathsf{V}}$, then either $\mathrm{a} = 0_{\mathsf{F}}$ or $\mathrm{v} = 0_{\mathsf{V}}$.

Proof Parts (i) and (ii) follow in the same manner as part (i) of Proposition 4.1.6.

(iii) For some $v \in \mathsf{V}$ we compute

$$av = a(v + 0_{\mathsf{V}}) = av + a0_{\mathsf{V}}.$$

Therefore,

$$av + (-(av)) = av + (-(av)) + a0_{\mathsf{V}} \quad\implies\quad 0_{\mathsf{V}} = 0_{\mathsf{V}} + a0_{\mathsf{V}} = a0_{\mathsf{V}},$$

which gives the result.

(iv) For some $a \in \mathsf{F}$ we compute

$$av = (a + 0_{\mathsf{F}})v = av + 0_{\mathsf{F}}v.$$

Therefore,

$$av + (-(av)) = av + (-(av)) + 0_{\mathsf{F}}v \quad \Longrightarrow \quad 0_{\mathsf{V}} = 0_{\mathsf{V}} + 0_{\mathsf{F}}v = 0_{\mathsf{F}}v,$$

giving the result.

(v) We have

$$0_{\mathsf{V}} = a0_{\mathsf{V}} = a(v + (-v)) = av + a(-v).$$

Therefore, $a(-v) = -(av)$. Similarly,

$$0_{\mathsf{V}} = 0_{\mathsf{F}}v = (a - a)v = av + (-a)v.$$

Therefore $(-a)v = -(av)$.

(vi) Suppose that $av = 0_{\mathsf{V}}$. If $a = 0_{\mathsf{F}}$ then there is nothing to prove. If $a \neq 0_{\mathsf{F}}$ then we have

$$0_{\mathsf{V}} = a^{-1}0_{\mathsf{V}} = a^{-1}(av) = (a^{-1}a)v = 1_{\mathsf{F}}v = v,$$

which gives the result. ■

In this section it will be convenient to have on hand the notion of a homomorphism of vector spaces. This is a topic about which we will have much to say in Chapter **??**, but here we simply give the definition.

4.3.4 Definition (Linear map) Let F be a field and let U and V be F-vector spaces. An ***F-homomorphism*** of U and V, or equivalently an ***F-linear map*** between U and V, is a map $\mathsf{L}\colon \mathsf{U} \to \mathsf{V}$ having the properties that

(i) $\mathsf{L}(u_1 + u_2) = \mathsf{L}(u_1) + \mathsf{L}(u_2)$ for every $u_1, u_2 \in \mathsf{U}$ and

(ii) $\mathsf{L}(au) = a\mathsf{L}(u)$ for every $a \in \mathsf{F}$ and $u \in \mathsf{U}$.

An F-homomorphism L is an ***F-monomorphism*** (resp. ***F-epimorphism***, ***F-isomorphism***) if L is injective (resp. surjective, bijective). If there exists an isomorphism between F-vector spaces U and V, then U and V are ***F-isomorphic***. An F-homomorphism from V to itself is called an ***F-endomorphismmissing stuff*** of V. The set of F-homomorphisms from U to V is denoted by $\mathrm{Hom}_{\mathsf{F}}(\mathsf{U}; \mathsf{V})$, and the set of F-endomorphisms of V is denoted by $\mathrm{End}_{\mathsf{F}}(\mathsf{V})$. •

We shall frequently simply call an "F-homomorphism" or an "F-linear map " a "homomorphism" or a "linear map" when F is understood. We postpone to Section **??** an exposition of the properties of linear maps, as well as a collection of illustrative examples. In this section we shall principally encounter a few examples of isomorphisms.

4.3.2 Subspaces

As with most algebraic objects, with vector spaces it is interesting to talk about subsets that respect the structure.

4.3.5 Definition Let F be a field. A nonempty subset U of an F-vector space V is a ***vector subspace***, or simply a ***subspace***, if $u_1 + u_2 \in \mathsf{U}$ for all $u_1, u_2 \in \mathsf{U}$ and if $au \in \mathsf{U}$ for all $a \in \mathbb{F}$ and all $u \in \mathsf{U}$. •

As we saw with subgroups and subrings, subspaces are themselves vector spaces.

4.3.6 Proposition (A vector subspace is a vector space) *Let* F *be a field. A nonempty subset* $\mathsf{U} \subseteq \mathsf{V}$ *of an* F*-vector space* V *is a subspace if and only if* U *is a vector space using the operations of vector addition and scalar multiplication in* V*, restricted to* U*.*

Proof This is Exercise 4.3.11. ■

Let us give some examples of subspaces. We leave the straightforward verifications of our claims as exercises.

4.3.7 Examples (Subspaces)

1. For each $n \in \mathbb{Z}_{>0}$, F^n can be regarded as a subspace of F_0^∞ by tacking on zeros to the n-tuple in F^n to get a sequence indexed by $\mathbb{Z}_{>0}$.
2. The subset F_0^∞ of F^∞ is a subspace.
3. For each $k \in \mathbb{Z}_{\geq 0}$, $\mathsf{F}_k[\xi]$ is a subspace of $\mathsf{F}[\xi]$. However, the set of polynomials of degree k is *not* a subspace of $\mathsf{F}[\xi]$. Why?
4. In Exercise 4.3.10 the reader can verify that, for $r \in \mathbb{Z}_{>0}$, the set $\mathsf{C}^r(I;\mathbb{R})$ of r-times continuously differentiable $\mathbb{R}$-valued functions defined on an interval I is a $\mathbb{R}$-vector space. In fact, it is a subspace of $\mathsf{C}^0(I;\mathbb{R})$. •

Analogously with homomorphisms of groups and rings, there are two natural subspaces associated with a homomorphism of vector spaces.

4.3.8 Definition (Kernel and image of linear map) Let F be a vector space, let U and V be F-vector spaces, and let $\mathsf{L} \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{U};\mathsf{V})$.

(i) The ***image*** of L is $\mathrm{image}(\mathsf{L}) = \{\mathsf{L}(u) \mid u \in \mathsf{U}\}$.

(ii) The ***kernel*** of L is $\ker(\mathsf{L}) = \{u \in \mathsf{U} \mid \mathsf{L}(u) = 0_{\mathsf{V}}\}$. •

It is straightforward to verify that the image and kernel are subspaces.

4.3.9 Proposition (Kernel and image are subspaces) *Let* F *be a field, let* U *and* V *be* F*-vector spaces, and let* $\mathsf{L} \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{U};\mathsf{V})$*. Then* $\mathrm{image}(\mathsf{L})$ *and* $\ker(\mathsf{L})$ *are subspaces of* V *and* U*, respectively.*

Proof This is Exercise 4.3.16. ■

An important sort of subspace arises from taking sums of vectors with arbitrary coefficients in the field over which the vector space is defined. To make this more formal, we have the following definition.

4.3.10 Definition (Linear combination) Let F be a field and let V be an F-vector space. If $S \subseteq \mathsf{V}$ is nonempty, a ***linear combination*** from S is an element of V of the form

$$c_1 v_1 + \cdots + c_k v_k,$$

where $c_1, \dots, c_k \in \mathsf{F}$ and $v_1, \dots, v_k \in S$. We call $c_1, \dots, c_k$ the ***coefficients*** in the linear combination. •

The important feature of the set of linear combinations from a subset of a vector space is that they form a subspace.

4.3.11 Proposition (The set of linear combinations is a subspace) *If* F *is a field, if* V *is an* F*-vector space, and if* $\mathsf{S} \subseteq \mathsf{V}$ *is nonempty, then the set of linear combinations from* S *is a subspace of* V*. Moreover, this subspace is the smallest subspace of* V *containing* S*.*

Proof Let

$$B = b_1u_1 + \cdots + b_lv_l, \quad C = c_1v_1 + \cdots + c_kv_k$$

be linear combinations from S and let $a \in \mathsf{F}$. Then

$$B + C = b_1u_1 + \cdots + b_lu_l + c_1v_1 + \cdots + c_kv_k$$

is immediately a linear combination from S with vectors $u_1, \dots, u_l, v_1, \dots, v_k$ and coefficients $b_1, \dots, b_l, c_1, \dots, c_k$. Also

$$aC = (ac_1)v_1 + \cdots + (ac_k)v_k$$

is a linear combination from S with vectors $v_1, \dots, v_k$ and coefficients $ac_1, \dots, ac_k$. Thus $B + C$ and aC are linear combinations from S.

Now let U be a subspace of V containing S. If $c_1v_1 + \cdots + c_kv_k$ is a linear combination from S then, since $S \subseteq \mathsf{U}$ and since U is a subspace, $c_1v_1 + \cdots + c_kv_k \in \mathsf{U}$. Therefore, U contains the set of linear combinations from S, and hence follows the second assertion of the proposition. ■

Based on the preceding result we have the following definition. Note that the definition is "geometric," whereas the proposition gives a more concrete version in that the explicit form of elements of the subspace are given.

4.3.12 Definition (Subspace generated by a set) If F is a field, if V is an F-vector space, and if $S \subseteq \mathsf{V}$ is nonempty, then the ***subspace generated by* S** is the smallest subspace of V containing S. This subspace is denoted by $\mathrm{span}_{\mathsf{F}}(S)$. •

We close this section with a definition of a "shifted subspace" which will come up in our discussion in Sections **??** and **??**.

4.3.13 Definition (Affine subspace) Let F be a field and let V be an F-vector space. A subset $\mathsf{A} \subseteq \mathsf{V}$ is an ***affine subspace*** if there exists $v_0 \in \mathsf{V}$ and a subspace U of V such that

$$\mathsf{A} = \{v_0 + u \mid u \in \mathsf{U}\}.$$

The subspace U is the ***linear part*** of A. •

Intuitively, an affine subspace is a subspace U shifted by the vector v_0. Let us give some simple examples of affine subspaces.

4.3.14 Examples (Affine subspaces)

1. Every subspace is also an affine subspace "shifted" by the zero vector.
2. If U is a subspace of a vector space V and if $u_0 \in \mathsf{U}$, then the affine subspace

$$\{u_0 + u \mid u \in \mathsf{U}\}$$

is simply the subspace U. That is to say, if we shift a subspace by an element of itself, the affine subspace is simply a subspace.

3. Let $\mathsf{V} = \mathbb{R}^2$. The vertical line

$$\{(1,0) + (0,y) \mid y \in \mathbb{R}\}$$

through the point $(1,0)$ is an affine subspace. •

4.3.3 Linear independence

The notion of linear independence lies at the heart of understanding much of the theory of vectors spaces, and the associated topic of linear algebra which we treat in detail in Chapter **??**. The precise definition we give for linear independence is one that can be difficult to understand on a first encounter. However, it is important to understand that this definition has, in actuality, been carefully crafted to be maximally useful; the definition in its precise form is used again and again in proofs in this section and in Chapter **??**.

4.3.15 Definition (Linearly independent) Let F be a field and let V be an F-vector space.

(i) A finite family $(v_1, \dots, v_k)$ of vectors in V is ***linearly independent*** if the equality

$$c_1 v_1 + \dots + c_k v_k = 0_{\mathsf{V}}, \qquad c_1, \dots, c_k \in \mathsf{F},$$

is satisfied only if $c_1 = \dots = c_k = 0_{\mathsf{F}}$.

(ii) A finite set $S = \{x_j \mid j \in \{1, \dots, k\}\}$ is linearly independent if the finite family corresponding to the set is linearly independent.

(iii) An nonempty family $(v_a)_{a \in A}$ of vectors in V is ***linearly independent*** if every finite subfamily of $(v_a)_{a \in A}$ is linearly independent.

(iv) A nonempty subset $S \subseteq \mathsf{V}$ is ***linearly independent*** if every nonempty finite subset of S is linearly independent.

(v) A nonempty family $(v_a)_{a \in A}$ if vectors in V is ***linearly dependent*** if it is not linearly independent.

(vi) A nonempty subset $S \subseteq \mathsf{V}$ is ***linearly dependent*** if it is not linearly independent. •

The definition we give is not quite the usual one since we define linear independence and linear dependence for both sets of vectors and families of vectors. Corresponding to any set $S \subseteq \mathsf{V}$ of vectors is a family of vectors in a natural way: $(v)_{v \in S}$. Thus one can, in actuality, get away with only defining linear independence and linear dependence for families of vectors. However, since most references will consider sets of vectors, we give both flavours of the definition. Let us see with a simple example that only dealing with sets of vectors may not suffice.

4.3.16 Example (Sets of vectors versus families of vectors) Let F be a field and let $\mathsf{V} = \mathsf{F}^2$. Define $v_1 = (1_{\mathsf{F}}, 0_{\mathsf{F}})$ and $v_2 = (1_{\mathsf{F}}, 0_{\mathsf{F}})$. Then the family (v_1, v_2) is linearly dependent since $1_{\mathsf{F}} v_1 - 1_{\mathsf{F}} v_2 = 0_{\mathsf{V}}$. However, since $\{v_1, v_2\} = \{(1_{\mathsf{F}}, 0_{\mathsf{F}}\}$, this set is, in fact, linearly independent. •

As can easily be gleaned from this example, the distinction between linearly independent sets and linearly independent families only arises when the family contains the same vector in two places. We shall frequently talk about sets rather than families, accepting that in doing so we disallow the possibility of considering that two vectors in the set might be the same.

There is a potential inconsistency with the above definition of a general linearly independent set. Specifically, if $S = (v_1, \dots, v_k)$ is a finite family of vectors, then Definition 4.3.15 proposes two definitions of linear independence, one from part (i) and one from part (iv). To resolve this we prove the following result.

4.3.17 Proposition (Subsets of finite linearly independent sets are linearly independent) *Let* F *be a field, let* V *be an* F*-vector space, and let* $(\mathrm{v}_1, \dots, \mathrm{v}_k)$ *be linearly independent according to part (i) of Definition 4.3.15. Then any nonempty subfamily of* $(\mathrm{v}_1, \dots, \mathrm{v}_k)$ *is linearly independent.*

Proof Let $(v_{j_1}, \dots, v_{j_l})$ be a nonempty subfamily of $(v_1, \dots, v_k)$ and suppose that

$$c_1 v_{j_1} + \dots + c_l v_{j_l} = 0_{\mathsf{V}}.$$

Let $\{j_{l+1}, \dots, j_k\}$ be a distinct set of indices for which $\{1, \dots, k\} = \{j_1, \dots, j_l, j_{l+1}, j_k\}$. Then

$$c_1 v_{j_1} + \dots + c_l v_{j_l} + 0_{\mathsf{F}} v_{j_{l+1}} + \dots + 0_{\mathsf{F}} v_{j_k} = 0_{\mathsf{V}}.$$

Since the set $(v_1, \dots, v_k)$ is linearly independent, it follows that $c_1 = \dots = c_l = 0_{\mathsf{F}}$, giving the result. ■

Let us give some examples of linearly independent and linearly dependent sets to illustrate the ideas.

4.3.18 Examples (Linear independence)

1. In the F-vector space F^n consider the n vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ defined by

$$\boldsymbol{e}_j = (0, \dots, 0, \underbrace{1_{\mathsf{F}}}_{j\text{th position}}, 0, \dots, 0).$$

We claim that these vectors are linearly independent. Indeed, suppose that

$$c_1 \boldsymbol{e}_1 + \dots + c_n \boldsymbol{e}_n = 0_{\mathsf{F}^n}$$

for $c_1, \dots, c_n \in \mathsf{F}$. Using the definition of vector addition and scalar multiplication in F^n this means that

$$(c_1, \dots, c_n) = (0, \dots, 0),$$

which immediately gives $c_1 = \dots = c_n = 0_{\mathsf{F}}$. This gives linear independence, as desired.

2. In the F-vector space F_0^∞ define vectors $\boldsymbol{e}_j$, $j \in \mathbb{Z}_{>0}$, by asking that $\boldsymbol{e}_j$ be the sequence consisting of zeros except for the jth term in the sequence, which is 1_{F}.

We claim that the family $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ is linearly independent. Indeed. let $\boldsymbol{e}_{j_1},\dots,\boldsymbol{e}_{j_k}$ be a finite subset of $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$. Then suppose that

$$c_1\boldsymbol{e}_{j_1}+\dots+c_k\boldsymbol{e}_{j_k}=0_{\mathsf{F}_0^\infty}$$

for $c_1,\dots,c_k\in\mathsf{F}$. Using the definition of vector addition and scalar multiplication in F_0^∞, the linear combination $c_1\boldsymbol{e}_{j_1}+\dots+c_k\boldsymbol{e}_{j_k}$ is equal to the sequence $(a_l)_{j\in\mathbb{Z}_{>0}}$ in F given by

$$a_l=\begin{cases}c_r, & l=j_r \text{ for some } r\in\{1,\dots,k\},\\ 0_{\mathsf{F}}, & \text{otherwise.}\end{cases}$$

Clearly this sequence is equal to zero if and only if $c_1=\dots=c_k=0_{\mathsf{F}}$, thus showing that $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ is linearly independent.

3. Since F_0^∞ is a subspace of F^∞, it follows easily that the family $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ is linearly independent in F^∞.***missing stuff***
4. In the F-vector space $\mathsf{F}_k[\xi]$ of polynomials of degree at most k the family $(1,\xi,\dots,\xi^k)$ is linearly independent. Indeed, suppose that

$$c_0+c_1\xi+\dots+c_k\xi^k=0_{\mathsf{F}[\xi]} \tag{4.1}$$

for $c_0,c_1,\dots,c_k\in\mathsf{F}$. One should now recall the definition of $\mathsf{F}[\xi]$ as sequences in F for which a finite number of elements in the sequence are nonzero. The elements in the sequence, recall, are simply the coefficients of the polynomial. Therefore, a polynomial is the zero polynomial if and only if all of its coefficients are zero. In particular, (4.1) holds if and only if $c_0=c_1=\dots=c_k=0_{\mathsf{F}}$.

5. In the vector space $\mathsf{F}[\xi]$ we claim that the set $(\xi^j)_{j\in\mathbb{Z}_{\geq0}}$ is linearly independent. To see why this is so, choose a finite subfamily $(\xi^{j_1},\dots,\xi^{j_k})$ from the family $(\xi^j)_{j\in\mathbb{Z}_{\geq0}}$ and suppose that

$$c_1\xi^{j_1}+\dots+c_k\xi^{j_k}=0_{\mathsf{F}[\xi]} \tag{4.2}$$

for some $c_1,\dots,c_k\in\mathsf{F}$. As we argued in the previous example, a polynomial is zero if and only if all of its coefficients is zero. Therefore, (4.2) holds if and only if $c_1=\dots=c_k=0_{\mathsf{F}}$, thus showing linear independence of the family $(\xi^j)_{j\in\mathbb{Z}_{\geq0}}$.

6. In the $\mathbb{R}$-vector space $\mathsf{C}^0([0,\pi];\mathbb{R})$ define vectors (i.e., functions) $\cos_j\colon I\to\mathbb{R}$, $j\in\mathbb{Z}_{\geq0}$, and $\sin_j\colon I\to\mathbb{R}$, $j\in\mathbb{Z}_{>0}$, by

$$\cos_j=\cos(jx),\quad \sin_j(x)=\sin(jx).$$

We claim that the family $(\cos_j)_{j\in\mathbb{Z}_{\geq0}}\cup(\sin_j)_{j\in\mathbb{Z}_{>0}}$ is linearly independent. To see this, suppose that a finite linear combination of these vectors vanishes:

$$a_1\cos_{j_1}+\dots+a_l\cos_{j_l}+b_1\sin_{k_1}+\dots+b_m\sin_{k_m}=0_{\mathsf{C}^0([0,2\pi];\mathbb{R})}, \tag{4.3}$$

for $a_1, \dots, a_l, b_1, \dots, b_m \in \mathbb{R}$. Now multiply (4.3) by the function $\cos_{j_r}$ for some $r \in \{1, \dots, l\}$ and integrate both sides of the equation over the interval $[0, 2\pi]$:

$$a_1 \int_0^{2\pi} \cos_{j_1}(x)\cos_{j_r}(x)\,\mathrm{d}x + \dots + a_l \int_0^{2\pi} \cos_{j_l}(x)\cos_{j_r}(x)\,\mathrm{d}x \\ + b_1 \int_0^{2\pi} \sin_{k_1}(x)\cos_{j_r}(x)\,\mathrm{d}x + \dots + b_m \int_0^{2\pi} \sin_{k_m}(x)\cos_{j_r}(x)\,\mathrm{d}x = 0. \quad (4.4)$$

Now we recall the following trigonometric identities

$$\cos(a)\cos(b) = \tfrac{1}{2}(\cos(a-b) + \cos(a+b)), \quad \cos(a)\sin(b) = \tfrac{1}{2}(\sin(a+b) - \sin(a-b)), \\ \sin(a)\sin(b) = \tfrac{1}{2}(\cos(a-b) - \cos(a+b)), \\ \cos^2(a) = \tfrac{1}{2}(1 + \cos(2a)), \quad \sin^2(a) = \tfrac{1}{2}(1 - \cos(2a),$$

for $a, b \in \mathbb{R}$. The above identities are easily proved using Euler's formula $\mathrm{e}^{\mathrm{i}x} = \cos(x) + \mathrm{i}\sin(x)$ and properties of the exponential function. We recommend that the reader learn these derivations and then overwrite that portion of their memory used for storing trigonometric identities with something useful like, say, sports statistics or lines from their favourite movies. The above trigonometric identities can now be used, along with the derivative (and hence integral, by the Fundamental Theorem of Calculus) rules for trigonometric functions to derive the following identities for $j, k \in \mathbb{Z}_{>0}$:

$$\int_0^{2\pi} \cos(jx)\cos(kx)\,\mathrm{d}x = \begin{cases} 0, & j \neq k, \\ \pi, & j = k, \end{cases} \\ \int_0^{2\pi} \cos(jx)\sin(kx)\,\mathrm{d}x = 0, \\ \int_0^{2\pi} \sin(jx)\sin(kx)\,\mathrm{d}x = \begin{cases} 0, & j \neq k, \\ \pi, & j = k, \end{cases} \\ \int_0^{2\pi} \cos(0x)\cos(0x)\,\mathrm{d}x = 2\pi, \\ \int_0^{2\pi} \cos(0x)\cos(kx)\,\mathrm{d}x = 0, \\ \int_0^{2\pi} \cos(0x)\sin(kx)\,\mathrm{d}x = 0.$$

Applying these identities to (4.4) gives $\pi a_r = 0$ if $j_r \neq 0$ and gives $2\pi a_r = 0$ if $j_r = 0$. In either case we deduce that $a_r = 0$, $r \in \{1, \dots, l\}$. In like manner, multiplying (4.3) by $\sin_{k_s}$, $s \in \{1, \dots, m\}$, and integrating over the interval $[0, 2\pi]$ gives $b_s = 0$, $s \in \{1, \dots, m\}$. This shows that the coefficients in the linear combination (4.3) are zero, and, therefore, that the set $(\cos_j)_{j \in \mathbb{Z}_{\geq 0}} \cup (\sin_j)_{j \in \mathbb{Z}_{>0}}$ is indeed linearly independent. •

The reader will hopefully have noticed strong similarities between Examples 1 and 4 and between Examples 2 and 5. This is not an accident, but is due to the fact that the vector spaces F^{k+1} and $\mathsf{F}_k[\xi]$ are isomorphic and that the vector spaces F_0^∞ and $\mathsf{F}[\xi]$ are isomorphic. The reader is asked to explicitly write isomorphisms of these vector spaces in Exercise 4.3.21.

Let us now prove some facts about linearly independent and linearly dependent sets.

4.3.19 Proposition (Properties of linearly (in)dependent sets) *Let* F *be a field, let* V *be an* F*-vector space, and let* $\mathsf{S} \subseteq \mathsf{V}$ *be nonempty. Then the following statements hold:*

(i) if $\mathsf{S} = \{\mathrm{v}\}$ *for some* $\mathrm{v} \in \mathsf{V}$*, then* S *is linearly independent if and only if* $\mathrm{v} \neq 0_{\mathsf{V}}$*;*

(ii) if $0_{\mathsf{V}} \in \mathsf{S}$ *then* S *is linearly dependent;*

(iii) if S *is linearly independent and if* $\mathrm{T} \subseteq \mathsf{S}$ *is nonempty, then* T *is linearly independent;*

(iv) if S *is linearly dependent and if* $\mathrm{T} \subseteq \mathsf{V}$*, then* $\mathsf{S} \cup \mathrm{T}$ *is linearly dependent;*

(v) if S *is linearly independent, if* $\{\mathrm{v}_1, \dots, \mathrm{v}_k\} \subseteq \mathsf{S}$*, and if*

$$\mathrm{a}_1\mathrm{v}_1 + \cdots + \mathrm{a}_k\mathrm{v}_k = \mathrm{b}_1\mathrm{v}_1 + \cdots + \mathrm{b}_k\mathrm{v}_k$$

for $\mathrm{a}_1, \dots, \mathrm{a}_k, \mathrm{b}_1, \dots, \mathrm{b}_k \in \mathsf{F}$*, then* $\mathrm{a}_j = \mathrm{b}_j$*,* $\mathrm{j} \in \{1, \dots, \mathrm{k}\}$*;*

(vi) if S *is linearly independent and if* $\mathrm{v} \notin \mathrm{span}_{\mathsf{F}}(\mathsf{S})$*, then* $\mathsf{S} \cup \{\mathrm{v}\}$ *is linearly independent.*

Proof (i) Note that $c0_{\mathsf{V}} = 0_{\mathsf{V}}$ if and only if $c = 0_{\mathsf{F}}$ by Proposition 4.3.3(vi). This is exactly equivalent to what we are trying to prove.

(ii) If $0_{\mathsf{V}} \in S$ then the finite subset $\{0_{\mathsf{V}}\}$ is linearly dependent by part (i).

(iii) Let $\{v_1, \dots, v_k\} \subseteq T \subseteq S$ and suppose that

$$c_1v_1 + \dots c_kv_k = 0_{\mathsf{V}}$$

for $c_1, \dots, c_k \in \mathsf{F}$. Since $\{v_1, \dots, v_k\} \subseteq S$ and since S is linearly independent, it follows that $c_1 = \cdots = c_k = 0_{\mathsf{F}}$.

(iv) Since S is linearly dependent there exists vectors $\{v_1, \dots, v_k\} \subseteq S$ and $c_1, \dots, c_k \in \mathsf{F}$ not all zero such that

$$c_1v_1 + \cdots + c_kv_k = 0_{\mathsf{V}}.$$

Since $\{v_1, \dots, v_k\} \subseteq S \cup T$, it follows that $S \cup T$ is linearly dependent.

(v) If

$$a_1v_1 + \cdots + a_kv_k = b_1v_1 + \cdots + a_kv_k,$$

then

$$(a_1 - b_1)v_1 + \cdots + (a_k - b_k)v_k = 0_{\mathsf{V}}.$$

Since the set $\{v_1, \dots, v_k\}$ is linearly independent, it follows that $a_j - b_j = 0_{\mathsf{F}}$ for $j \in \{1, \dots, k\}$, which gives the result.

(vi) Let $\{v_1, \dots, v_k\} \subseteq S \cup \{v\}$. If $\{v_1, \dots, v_k\} \subseteq S$ then the set is immediately linearly independent. If $\{v_1, \dots, v_k\} \not\subseteq S$, then we may without loss of generality suppose that $v_k = v$. Suppose that

$$c_1v_1 + \cdots + c_{k-1}v_{k-1} + c_kv_k = 0_{\mathsf{V}}.$$

First suppose that $c_k \neq 0_{\mathsf{F}}$. Then

$$v_k = -c_k^{-1}c_1v_1 + \cdots + c_k^{-1}c_{k-1}v_{k-1},$$

which contradicts the fact that $v_k \notin \operatorname{span}_{\mathsf{F}}(S)$. Thus we must have $c_k = 0_{\mathsf{F}}$. However, since S is linearly independent, it immediately follows that $c_1 = \cdots = c_{k-1} = 0_{\mathsf{F}}$. Thus $S \cup \{v\}$ is linearly independent. ■

4.3.4 Basis and dimension

The notion of the dimension of a vector space, which is derived from the concept of a basis, is an important one. Of particular importance is the dichotomy between vector spaces whose dimension is finite and those whose dimension is infinite. Essentially, finite-dimensional vector spaces, particularly those defined over $\mathbb{R}$, behave in a manner which often correspond somehow to our intuition. In infinite dimensions, however, our intuition can often lead us astray. And in these volumes we will be often interested in infinite-dimensional vector spaces. This infinite-dimensional case is complicated, and any sort of understanding will require understanding much of Chapter **??**.

For now, we get the ball rolling by introducing the idea of a basis.

4.3.20 Definition (Basis for a vector space) Let F be a field and let V be a vector space over F. A ***basis*** for V is a subset $\mathscr{B}$ of V with the properties that

(i) $\mathscr{B}$ is linearly independent and

(ii) $\operatorname{span}_{\mathsf{F}}(\mathscr{B}) = \mathsf{V}$. •

4.3.21 Remark (Hamel[1] basis) Readers who have had a first course in linear algebra should be sure to note that we do not require a basis to be a finite set. Nonetheless, the definition we give is probably exactly the same as the one encountered in a typical first course. What is different is that we have defined the notion of linear independence and the notion associated with the symbol "$\operatorname{span}_{\mathsf{F}}(\cdot)$" in a general way. Sometimes the word "basis" is reserved for finite sets of vectors, with the notion we give being called a ***Hamel basis***. •

Let us first prove that every vector space possesses a basis in the sense that we have defined the notion.

4.3.22 Theorem (Every vector space possesses a basis) *If F is a field and if V is an F-vector space, then there exists a basis for V.*

Proof Let $\mathscr{C}$ be the collection of subsets of V that are linearly independent. Such collections exist since, for example, $\{v\} \in \mathscr{C}$ if $v \in \mathsf{V}$ is nonzero. Place a partial order $\preceq$ on $\mathscr{C}$ by asking that $S_1 \preceq S_2$ if $S_1 \subseteq S_2$. Let $\mathscr{S} \subseteq \mathscr{C}$ be a totally ordered subset. Note that $\cup_{S\in\mathscr{S}} S$ is an element of $\mathscr{C}$. Indeed, let $\{v_1, \dots, v_k\} \subseteq \cup_{S\in\mathscr{S}} S$. Then $v_j \in S_j$ for some $S_j \in \mathscr{S}$. Let $j_0 \in \{1, \dots, k\}$ be chosen such that S_{j_0} is the largest of the sets $S_1, \dots, S_k$ according to the partial order $\preceq$, this being possible since $\mathscr{S}$ is totally ordered. Then $\{v_1, \dots, v_k\} \subseteq S_{j_0}$ and so $\{v_1, \dots, v_k\}$ is linearly independent since S_{j_0} is linearly independent. It is also evident that $\cup_{S\in\mathscr{S}} S$ is an upper bound for $\mathscr{S}$. Thus every totally ordered subset of $\mathscr{C}$ possesses an upper bound, and so by Zorn's Lemma

[1] Georg Karl Wilhelm Hamel (1877–1954) was a German mathematician whose contributions to mathematics were in the areas of function theory, mechanics, and the foundations of mathematics

possesses a maximal element. Let $\mathscr{B}$ be such a maximal element. By construction $\mathscr{B}$ is linearly independent. Let $v \in \mathsf{V}$ and suppose that $v \notin \operatorname{span}_{\mathsf{F}}(\mathscr{B})$. Then by Proposition 4.3.19(vi), $\mathscr{B} \cup \{v\}$ is linearly independent and $\mathscr{B} \subseteq \mathscr{B}\{v\}$. This contradicts the fact that $\mathscr{B}$ is maximal, and so it must hold that if $v \in \mathsf{V}$, then $v \in \operatorname{span}_{\mathsf{F}}(\mathscr{B})$. That is to say, $\operatorname{span}_{\mathsf{F}}(\mathscr{B}) = \mathsf{V}$. ■

One of the important properties of a basis is the following result.

4.3.23 Proposition (Unique representation of vectors in bases) *If* F *is a field, if* V *is an* F*-vector space, and if* $\mathscr{B}$ *is a basis for* V*, then, for* $\mathrm{v} \in \mathsf{V}$ *there exists a unique finite subset* $\{\mathrm{v}_1, \dots, \mathrm{v}_k\} \subseteq \mathscr{B}$ *and unique nonzero coefficients* $\mathrm{c}_1, \dots, \mathrm{c}_k \in \mathsf{F}$ *such that*

$$\mathrm{v} = \mathrm{c}_1\mathrm{v}_1 + \cdots + \mathrm{c}_k\mathrm{v}_k.$$

Proof Let $v \in \mathsf{V}$. Since $\operatorname{span}_{\mathsf{F}}(\mathscr{B}) = \mathsf{V}$, there exists $\{u_1, \dots, u_l\} \subseteq \mathscr{B}$ and $a_1, \dots, a_l \in \mathsf{F}$ such that

$$v = a_1u_1 + \cdots + a_lu_l. \tag{4.5}$$

Moreover, given the vectors $\{u_1, \dots, u_l\}$, the coefficients $a_1, \dots, a_l$ in (4.5) are unique. Let $\{v_1, \dots, v_k\} \subseteq \{u_1, \dots, u_l\}$ be these vectors for which the corresponding coefficient in (4.5) is nonzero. Denote by $c_1, \dots, c_k$ the coefficients in (4.5) corresponding to the vectors $\{v_1, \dots, v_k\}$. This gives the existence part of the result.

Suppose that $\{v'_1, \dots, v'_{k'}\} \subseteq \mathscr{B}$ and $c'_1, \dots, c'_{k'} \in \mathsf{F}^*$ satisfy

$$v = c'_1v'_1 + \cdots + c'_{k'}v'_{k'}.$$

Now take $\{w_1, \dots, w_m\}$ to be a set of vectors such that $\{w_1, \dots, w_m\} = \{v_1, \dots, v_k\} \cup \{v'_1, \dots, v'_{k'}\}$. Note that

$$\{v_1, \dots, v_k\}, \{v'_1, \dots, v'_{k'}\} \subseteq \{w_1, \dots, w_m\}.$$

Since $\{w_1, \dots, w_m\} \subseteq \mathscr{B}$ it is linearly independent. Therefore, by Proposition 4.3.19(v), there exists unique coefficients $b_1, \dots, b_m \in \mathsf{F}$ such that

$$v = b_1w_1 + \cdots + b_mw_m.$$

But we also have

$$v = c_1v_1 + \cdots + c_kv_k = c'_1v'_1 + \cdots + c'_{k'}v'_{k'}.$$

Therefore, it must hold that $\{v_1, \dots, v_k\} = \{v'_1, \dots, v'_{k'}\} = \{w_1, \dots, w_m\}$, and from this the result follows. ■

One of the more useful characterisations of bases is the following result.

4.3.24 Theorem (Linear maps are uniquely determined by their values on a basis) *Let* F *be a field, let* V *be an* F*-vector space, and let* $\mathscr{B} \subseteq \mathsf{V}$ *be a basis. Then, for any* F*-vector space* W *and any map* $\phi\colon \mathscr{B} \to \mathsf{W}$ *there exists a unique linear map* $\mathsf{L}_\phi \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{V}; \mathsf{W})$ *such that the diagram*

$$\begin{tikzcd}
\mathscr{B} \arrow[r, "\phi"] \arrow[d] & \mathsf{W} \\
\mathsf{V} \arrow[ur, "\mathsf{L}_\phi"'] &
\end{tikzcd}$$

commutes, where the vertical arrow is the inclusion.

Proof Denote $\mathscr{B} = \{e_i\}_{i \in I}$. If $v \in \mathsf{V}$ we have $v = \sum_{i \in I} v_i e_i$ for $v_i \in \mathsf{F}$, $i \in I$, all but finitely many of which are zero. Then define

$$\mathsf{L}_\phi(v) = \sum_{i \in I} v_i \phi(e_i).$$

This map is linear since

$$\mathsf{L}_\phi(u + v) = \sum_{i \in I} (u_i + v_i)\phi(e_i) = \sum_{i \in I} u_i \phi(e_i) + \sum_{i \in I} v_i \phi(e_i) = \mathsf{L}_\phi(u) + \mathsf{L}_\phi(v)$$

and

$$\mathsf{L}_\phi(av) = \sum_{i \in I} a v_i \phi(e_i) = a \sum_{i \in I} v_i \phi(e_i) = a\mathsf{L}_\phi(v),$$

where all manipulations make sense by virtue of the sums being finite. This gives the existence part of the theorem.

Suppose that $\mathsf{L} \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{V}; \mathsf{W})$ is another linear map for which the diagram in the theorem statement commutes. This implies that $\mathsf{L}(e_i) = \mathsf{L}_\phi(e_i)$ for $i \in I$. Now, if $v = \sum_{i \in I} v_i e_i$ is a finite linear combination of basis elements, then

$$\mathsf{L}\Big(\sum_{i \in I} v_i e_i\Big) = \sum_{i \in I} v_i \mathsf{L}(e_i) = \sum_{i \in I} v_i \mathsf{L}_\phi(e_i) = \mathsf{L}_\phi\Big(\sum_{i \in I} v_i e_i\Big),$$

giving $\mathsf{L} = \mathsf{L}_\phi$. ■

The theorem is very useful, and indeed often used, since it tells us that to define a linear map one need only define it on each vector of a basis.

As we shall shortly see, the notion of the dimension of a vector space relies completely on a certain property of any two bases for a vector space, namely that they have the same cardinality.

4.3.25 Theorem (Different bases have the same size) *If* F *is a field, if* V *is an* F*-vector space, and if* $\mathscr{B}_1$ *and* $\mathscr{B}_2$ *are two bases for* V*, then* $\mathrm{card}(\mathscr{B}_1) = \mathrm{card}(\mathscr{B}_2)$.

Proof The proof is broken into two parts, the first for the case when one of $\mathscr{B}_1$ and $\mathscr{B}_2$ is finite, and the second the case when both $\mathscr{B}_1$ and $\mathscr{B}_2$ are infinite.

Let us first prove the following lemma.

1 Lemma *If* $\{\mathrm{v}_1, \dots, \mathrm{v}_\mathrm{n}\}$ *is a basis for* V *then any set of* $\mathrm{n} + 1$ *vectors in* V *is linearly dependent.*

Proof We prove the lemma by induction on n. In the case when $n = 1$ we have $\mathsf{V} = \mathrm{span}_{\mathsf{F}}(v_1)$. Let $u_1, u_2 \in \mathsf{V}$ so that $u_1 = a_1 v_1$ and $u_2 = a_2 v_1$ for some $a_1, a_2 \in \mathsf{F}$. If either u_1 or u_2 is zero then the set $\{u_1, u_2\}$ is immediately linearly dependent by Proposition 4.3.19(ii). Thus we can assume that a_1 and a_2 are both nonzero. In this case we have

$$a_2 u_1 - a_1 u_2 = a_2(a_1 v_1) - a_1(a_2 v_1) = 0_{\mathsf{V}},$$

so that $\{u_1, u_2\}$ is not linearly independent. Now suppose that the lemma holds for $n \in \{1, \dots, k\}$ and let $\{v_1, \dots, v_{k+1}\}$ be a basis for V. Consider a set $\{u_1, \dots, u_{k+2}\}$ and write

$$u_s = \sum_{r=1}^{k+1} a_{rs} v_r, \qquad s \in \{1, \dots, k+2\}.$$

First suppose that $a_{1s} = 0_{\mathsf{F}}$ for all $s \in \{1, \dots, k+2\}$. It then holds that $\{u_1, \dots, u_{k+2}\} \subseteq \operatorname{span}_{\mathsf{F}}(v_2, \dots, v_{k+1})$. By the induction hypothesis, since $\operatorname{span}_{\mathsf{F}}(v_2, \dots, v_{k+1})$ has basis $\{v_2, \dots, v_{k+1}\}$, it follows that $\{u_1, \dots, u_{k+1}\}$ is linearly dependent, and so $\{u_1, \dots, u_{k+2}\}$ is also linearly dependent by Proposition 4.3.19(iv). Thus we suppose that not all of the coefficients a_{1s}, $s \in \{1, \dots, k+2\}$ is zero. For convenience, and without loss of generality, suppose that $a_{11} \neq 0_{\mathsf{F}}$. Then

$$a_{11}^{-1}u_1 = v_1 + a_{11}^{-1}a_{21}v_2 + \dots + a_{11}^{-1}a_{k+1,1}v_{k+1}.$$

We then have

$$u_s - a_{11}^{-1}a_{1s}u_1 = \sum_{r=2}^{k+1}(a_{rs} + a_{1s}a_{11}^{-1}a_{r1})v_r, \qquad s \in \{2, \dots, k+2\}.$$

meaning that $u_s - a_{11}^{-1}a_{1s}u_1 \in \operatorname{span}_{\mathsf{F}}(v_2, \dots, v_{k+1})$ for $s \in \{2, \dots, k+2\}$. By the induction hypothesis it follows that the set $\{u_2 - a_{11}^{-1}a_{12}u_1, \dots, u_{k+2} - a_{11}^{-1}a_{1,k+2}u_1\}$ is linearly dependent. We claim that this implies that $\{u_1, u_2, \dots, u_{k+2}\}$ is linearly dependent. Indeed, let $c_2, \dots, c_{k+2} \in \mathsf{F}$ be not all zero and such that

$$c_2(u_2 - a_{11}^{-1}a_{12}u_1) + \dots + c_{k+2}(u_{k+2} - a_{11}^{-1}a_{1,k+2}u_1) = 0_{\mathsf{V}}.$$

Then

$$(-c_2a_{11}^{-1}a_{12} - \dots - c_{k+2}a_{11}^{-1}a_{1,k+2})u_1 + c_2u_2 + \dots + c_{k+2}u_{k+2} = 0_{\mathsf{V}}.$$

Since not all of the coefficients $c_2, \dots, c_{k+2}$ are zero, it follows that $\{u_1, u_2, \dots, u_{k+2}\}$ is linearly dependent. This completes the proof. ▼

Now consider the case when either $\mathscr{B}_1$ or $\mathscr{B}_2$ is finite. Thus, without loss of generality suppose that $\mathscr{B}_1 = \{v_1, \dots, v_n\}$. It follows that $\mathscr{B}_2$ can have at most n elements. Thus $\mathscr{B}_2 = \{u_1, \dots, u_m\}$ for $m \leq n$. But, since $\mathscr{B}_2$ is a basis, it also holds that $\mathscr{B}_1$ must have at most m elements. Thus $n \leq m$, and so $m = n$ and thus $\operatorname{card}(\mathscr{B}_1) = \operatorname{card}(\mathscr{B}_2)$.

Now let us turn to the general case when either or both of $\mathscr{B}_1$ and $\mathscr{B}_2$ are infinite. For $u \in \mathscr{B}_1$ let $\mathscr{B}_2(u)$ be the unique finite subset $\{v_1, \dots, v_k\}$ of $\mathscr{B}_2$ such that

$$u = c_1v_1 + \dots + c_kv_k$$

for some $c_1, \dots, c_k \in \mathsf{F}^*$. We now prove a lemma.

2 Lemma *If* $\mathrm{v} \in \mathscr{B}_2$ *then there exists* $\mathrm{u} \in \mathscr{B}_1$ *such that* $\mathrm{v} \in \mathscr{B}_2(\mathrm{u})$*.*

Proof Suppose otherwise. Thus suppose that there exists $v \in \mathscr{B}_2$ such that, for every $u \in \mathscr{B}_1$, $v \notin \mathscr{B}_2(u)$. We claim that $\mathscr{B}_1 \cup \{v\}$ is then linearly independent. Indeed, let $\{v_1, \dots, v_k\} \subseteq \mathscr{B}_1 \cup \{v\}$. If $\{v_1, \dots, v_k\} \subseteq \mathscr{B}_1$ then we immediately have that $\{v_1, \dots, v_k\}$ is linearly independent. So suppose that $\{v_1, \dots, v_k\} \not\subseteq \mathscr{B}_1$, and suppose without loss of generality that $v_k = v$. Let $c_1, \dots, c_k \in \mathsf{F}$ satisfy

$$c_1v_1 + \dots + c_kv_k = 0_{\mathsf{V}}.$$

If $c_k \neq 0_{\mathsf{F}}$ then

$$v = -c_k^{-1}c_1v_1 + \dots - c_k^{-1}c_{k-1}v_{k-1},$$

implying that $v \in \operatorname{span}_{\mathsf{F}}(v_1, \dots, v_{k-1})$. We can thus write v as a linear combination of vectors from the finite subsets $\mathscr{B}_2(v_j)$, $j \in \{1, \dots, k-1\}$. Let $\{w_1, \dots, w_m\}$ be a set of distinct vectors with the property that

$$\{w_1, \dots, w_m\} = \cup_{j=1}^{k-1} \mathscr{B}_2(v_j).$$

Thus $\mathscr{B}_2(v_j) \subseteq \{w_1, \dots, w_m\}$ for $j \in \{1, \dots, k-1\}$. It then follows that $v \in \operatorname{span}_{\mathsf{F}}(w_1, \dots, w_m)$. However, since $v \notin \{w_1, \dots, w_m\}$ by our assumption that $v \notin \mathscr{B}_2(u)$ for every $u \in \mathscr{B}_1$, it follows that $\{v, w_1, \dots, w_m\}$ is linearly independent, which is a contradiction. Therefore, $c_k = 0_{\mathsf{F}}$.

On the other hand, if $c_k = 0_{\mathsf{F}}$ then it immediately follows that $c_1 = \dots = c_{k-1} = 0_{\mathsf{F}}$ since $\{v_1, \dots, v_{k-1}\} \subseteq \mathscr{B}_1$ and since $\mathscr{B}_1$ is linearly independent. Therefore, $\mathscr{B}_1 \cup \{v\}$ is indeed linearly independent. In particular, $v \notin \operatorname{span}_{\mathsf{F}}(\mathscr{B}_1)$, contradicting the fact that $\mathscr{B}_1$ is a basis. ▼

From the lemma we know that $\mathscr{B}_2 = \cup_{u \in \mathscr{B}_1} \mathscr{B}_2(u)$. By the definition of multiplication of cardinal numbers, and using the fact that $\operatorname{card}(\mathbb{Z}_{>0})$ exceeds every finite cardinal number, we have

$$\operatorname{card}(\mathscr{B}_2) \leq \operatorname{card}(\mathscr{B}_1) \operatorname{card}(\mathbb{Z}_{>0}).$$

By Corollary **??** it follows that $\operatorname{card}(\mathscr{B}_2) \leq \operatorname{card}(\mathscr{B}_1)$. By interchanging the rôles of $\mathscr{B}_1$ and $\mathscr{B}_2$ we can also show that $\operatorname{card}(\mathscr{B}_1) \leq \operatorname{card}(\mathscr{B}_2)$. By the Cantor–Schröder–Bernstein Theorem, $\operatorname{card}(\mathscr{B}_1) = \operatorname{card}(\mathscr{B}_2)$. ■

Let us give some other useful constructions concerning bases. The proofs we give are valid for arbitrary bases. We invite the reader to give proofs in the case of finite bases in Exercise 4.3.18.

4.3.26 Theorem (Bases and linear independence) *Let* F *be a field and let* V *be an* F*-vector space. For a subset* $\mathsf{S} \subseteq \mathsf{V}$*, the following statements hold:*

(i) if S *is linearly independent, then there exists a basis* $\mathscr{B}$ *for* V *such that* $\mathsf{S} \subseteq \mathscr{B}$*;*

(ii) if $\operatorname{span}_{\mathsf{F}}(\mathsf{S}) = \mathsf{V}$*, then there exists a basis* $\mathscr{B}$ *for* V *such that* $\mathscr{B} \subseteq \mathsf{S}$*.*

Proof (i) Let $\mathscr{C}(S)$ be the collection of linearly independent subsets of V which contain S. Since $S \in \mathscr{C}(S)$, $\mathscr{C}(S) \neq \emptyset$. The set $\mathscr{C}(S)$ can be partially ordered by inclusion. Thus $S_1 \preceq S_2$ if $S_1 \subseteq S_2$. Just as in the proof of Theorem 4.3.22, every totally ordered subset of $\mathscr{C}(S)$ has an upper bound, and so $\mathscr{C}(S)$ possesses a maximal element $\mathscr{B}$ by Zorn's Lemma. This set may then be shown to be a basis just as in the proof of Theorem 4.3.22.

(ii) Let $\mathscr{D}(S)$ be the collection of linearly independent subsets of S, and partially order $\mathscr{D}(S)$ by inclusion, just as we partially ordered $\mathscr{C}(S)$ in part (i). JUst as in the proof of Theorem 4.3.22, every totally ordered subset of $\mathscr{D}(S)$ has an upper bound, and so $\mathscr{D}(S)$ possesses a maximal element $\mathscr{B}$. We claim that every element of S is a linear combination of elements of $\mathscr{B}$. Indeed, if this were not the case, then there exists $v \in S$ such that $v \notin \operatorname{span}_{\mathsf{F}}(\mathscr{B})$. Then $\mathscr{B} \cup \{v\}$ is linear independent by Proposition 4.3.19(vi), and is also contained in S. This contradicts the maximality of $\mathscr{B}$, and so we indeed have $S \subseteq \operatorname{span}_{\mathsf{F}}(\mathscr{B})$. Therefore,

$$\operatorname{span}_{\mathsf{F}}(\mathscr{B}) = \operatorname{span}_{\mathsf{F}}(S) = \mathsf{V},$$

giving the theorem. ■

Now it makes sense to talk about the dimension of a vector space.

4.3.27 Definition (Dimension, finite-dimensional, infinite-dimensional) Let F be a field, let V be an F-vector space, and let $\mathscr{B}$ be a basis for V. The ***dimension*** of the vector space V, denoted by $\dim_{\mathsf{F}}(\mathsf{V})$, is the cardinal number $\mathrm{card}(\mathscr{B})$. If $\mathscr{B}$ is finite then V is ***finite-dimensional***, and otherwise V is ***infinite-dimensional***. We will slightly abuse notation and write $\dim_{\mathsf{F}}(\mathsf{V}) = \infty$ whenever V is infinite-dimensional. •

Let us give some examples of vector spaces of various dimensions.

4.3.28 Examples (Basis and dimension)

1. The trivial vector space $\mathsf{V} = \{0_{\mathsf{V}}\}$ consisting of the zero vector has $\emptyset$ as a basis.
2. The F-vector space F^n has as a basis the set $\mathscr{B} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ defined in Example 4.3.18–1. In that example, $\mathscr{B}$ was shown to be linearly independent. Also, since
$$(v_1, \dots, v_n) = v_1\boldsymbol{e}_1 + \dots + v_n\boldsymbol{e}_n,$$
it follows that $\mathrm{span}_{\mathsf{F}}(\mathscr{B}) = \mathsf{F}^n$. Thus $\dim_{\mathsf{F}}(\mathsf{F}^n) = n$. The basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is called the ***standard basis***.
3. The subspace F_0^∞ of F^∞ has a basis which is easily described. Indeed, it is easy to verify that $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ is a basis for F_0^∞. We adopt the notation from the finite-dimensional case and call this the ***standard basis***.
4. We next consider the F-vector space F^∞. Since $\mathsf{F}_0^\infty \subseteq \mathsf{F}^\infty$, and since the standard basis $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ is linearly independent in F^∞, we know by Theorem 4.3.26 that we can extend the standard basis for F_0^∞ to a basis for F^∞. This extension is nontrivial since, for example, the sequence $\{1_{\mathsf{F}}\}_{j\in\mathbb{Z}_{>0}}$ in F cannot be written as a finite linear combination of standard basis vectors. Thus the set $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}} \cup \{\{1_{\mathsf{F}}\}_{j\in\mathbb{Z}_{>0}}\}$ is linearly independent. This linearly set shares with the standard basis the property of being countable. It turns out, in fact, that any basis for F^∞ has the cardinality of $\mathbb{R}$, and so the process of tacking on linearly independent vectors to the standard basis for F_0^∞ will take a long time to produce a basis for F^∞. We will not understand this properly until Section **??**, where we will see that F^∞ is the algebraic dual of F_0^∞, and so thereby derive by general means the dimension of F^∞. For the moment we merely say that F^∞ is a much larger vector space than is F_0^∞.
5. In $\mathsf{F}_k[\xi]$, it is easy to verify that $\{1, \xi, \dots, \xi^k\}$ is a basis. Indeed, we have already shown that the set is linearly independent. It follows from the definition of $\mathsf{F}_k[\xi]$ that the set also generates $\mathsf{F}_k[\xi]$.
6. The set $\{\xi^j\}_{j\in\mathbb{Z}_{\geq 0}}$ forms a basis for $\mathsf{F}[\xi]$. Again, we have shown linear independence, and that this set generates $\mathsf{F}[\xi]$ follows by definition. •

4.3.29 Remark (Nonuniqueness of bases) Generally, it will not be the case that a vector spaces possesses a "natural" basis, although one might argue that the bases of Example 4.3.28 are fairly natural. But, even in cases where one might have a basis that is somehow distinguished, it is useful to keep in mind that other bases are possible, and that one should be careful not to rely overly on the comfort offered by a specific basis representation. In particular, if one is in the business of proving theorems using bases, one should make sure that what is being proved is independent of basis, if this is in fact what is intended. At this point in our presentation we do not have enough machinery at hand to explore this idea fully. However, we shall revisit this idea of basis independence in ***missing stuff***. Also, in ***missing stuff*** we shall discuss the matter of changing bases. •

Finally, let us prove the more or less obvious fact that dimension is preserved by isomorphism.

4.3.30 Proposition (Dimension characterises a vector space) *If* F *is a field and if* V_1 *and* V_2 *are* F*-vector spaces, then the following statements are equivalent:*

(i) V_1 *and* V_2 *are isomorphic;*

(ii) $\dim_{\mathsf{F}}(\mathsf{V}_1) = \dim_{\mathsf{F}}(\mathsf{V}_2)$*.*

Proof (i) $\Longrightarrow$ (ii) Let $\mathsf{L}\colon \mathsf{V}_1 \to \mathsf{V}_2$ be an isomorphism and let $\mathscr{B}_1$ be a basis for V_1. We claim that $\mathscr{B}_2 = \mathsf{L}(\mathscr{B}_1)$ is a basis for V_2. Let us first show that $\mathscr{B}_2$ is linearly independent. Let $v_1 = \mathsf{L}(u_1), \dots, v_k = \mathsf{L}(u_k) \in \mathscr{B}_2$ be distinct and suppose that

$$c_1 v_1 + \cdots = c_k v_k = 0_{\mathsf{V}_2}$$

for $c_1, \dots, c_k \in \mathsf{F}$. Since L is linear we have

$$\mathsf{L}(c_1 u_1 + \cdots + c_k u_k) = 0_{\mathsf{V}_2}.$$

Since L is injective, by Exercise 4.3.23 we have

$$c_1 u_1 + \cdots + c_k u_k = 0_{\mathsf{V}_1},$$

showing that $c_1 = \cdots = c_k = 0_{\mathsf{F}}$. Thus $\mathscr{B}_2$ is linearly independent. Moreover, for $v \in \mathsf{V}_2$ let $u = \mathsf{L}^{-1}(v)$ and then let $u_1, \dots, u_k \in \mathscr{B}_1$ and $c_1, \dots, c_k \in \mathsf{F}$ satisfy $u = c_1 u_1 + \cdots + c_k u_k$. Then

$$\mathsf{L}(u) = c_1 \mathsf{L}(u_1) + \cdots + c_k \mathsf{L}(u_k)$$

since L is linear. Therefore $v \in \operatorname{span}_{\mathsf{F}}(\mathscr{B}_2)$, and so $\mathscr{B}_2$ is indeed a basis. Since $\mathsf{L}|\mathscr{B}_1$ is a bijection onto $\mathscr{B}_2$ we have $\operatorname{card}(\mathscr{B}_2) = \operatorname{card}(\mathscr{B}_1)$, and this is the desired result.

(ii) $\Longrightarrow$ (i) Suppose that $\mathscr{B}_1$ and $\mathscr{B}_2$ are bases for V_1 and V_2, respectively, with the same cardinality. Thus there exists a bijection $\phi\colon \mathscr{B}_1 \to \mathscr{B}_2$. Now, by Theorem 4.3.24, define $\mathsf{L} \in \operatorname{Hom}_{\mathsf{F}}(\mathsf{V}_1; \mathsf{V}_2)$ by asking that $\mathsf{L}|\mathscr{B}_1 = \phi$. We claim that L is an isomorphism. To verify injectivity, suppose that $\mathsf{L}(u) = 0_{\mathsf{V}_2}$ for $u \in \mathsf{V}_1$. Write

$$u = c_1 u_1 + \cdots + c_k u_k$$

for $c_1, \dots, c_k \in \mathsf{F}$ and $u_1, \dots, u_k \in \mathscr{B}_1$. Then

$$0_{\mathsf{V}_2} = c_1 \mathsf{L}(u_1) + \cdots + c_k \mathsf{L}(u_k),$$

giving $c_j = 0_{\mathsf{F}}$, $j \in \{1, \dots, k\}$, since $\mathsf{L}(u_1), \dots, \mathsf{L}(u_k)$ are distinct elements of $\mathscr{B}_2$, and so linearly independent. Thus L is injective by Exercise 4.3.23. For surjectivity, let $v \in \mathsf{V}_2$ and write

$$v = c_1 v_1 + \cdots + c_k v_k$$

for $c_1, \dots, c_k \in \mathsf{F}$ and $v_1, \dots, v_k \in \mathscr{B}_2$. Then, if we define

$$u = c_1 \phi^{-1}(v_1) + \cdots + c_k \phi^{-1}(v_k) \in \mathsf{V}_2$$

we readily verify that $\mathsf{L}(u) = v$. ■

4.3.5 Intersections, sums, and products

In this section we investigate means of manipulating multiple subspaces and vector spaces. We begin by defining some constructions associated to subspaces of a vector space.

4.3.31 Definition (Sum and intersection) Let F be a field, let V be an F-vector space, and let $(\mathsf{U}_j)_{j\in J}$ be a family of subspaces of V indexed by a set J.

(i) The ***sum*** of $(\mathsf{U}_j)_{j\in J}$ is the subspace generated by $\cup_{j\in J}\mathsf{U}_j$, and is denoted by $\sum_{j\in J}\mathsf{U}_j$.

(ii) The ***intersection*** of $(\mathsf{U}_j)_{j\in J}$ is the set $\cap_{j\in J}\mathsf{U}_j$ (i.e., the set theoretic intersection). •

4.3.32 Notation (Finite sums of subspaces) If $\mathsf{U}_1,\dots,\mathsf{U}_k$ are a finite number of subspaces of an F-vector space V, then we will sometimes write

$$\sum_{j=1}^{k}\mathsf{U}_j=\mathsf{U}_1+\dots+\mathsf{U}_k. \qquad \bullet$$

4.3.33 Notation (Sum of subsets) We will also find it occasionally useful to be able to talk about sums of subsets that are not subspaces. Thus, if $(A_i)_{i\in I}$ is a family of subsets of an F-vector space V we denote by

$$\sum_{i\in I}A_i=\{v_{i_1}+\dots+v_{i_k}\mid i_1,\dots,i_k\in I\text{ distinct, }v_{i_j}\in A_{i_j},\ j\in\{1,\dots,k\},\ k\in\mathbb{Z}_{>0}\}.$$

Thus $\sum_{i\in I}A_i$ consists of finite sums of vectors from the subsets A_i, $i\in I$. Following our notation above, if $I=\{1,\dots,k\}$ then we write

$$\sum_{i\in I}A_i=A_1+\dots+A_k. \qquad \bullet$$

The sum and intersection are the subspace analogues of the set theoretic union and intersection, with the analogue being exact in the case of intersection. Note that the union of subspaces need not be a subspace (see Exercise 4.3.17). It is true that the intersection of subspaces is a subspace.

4.3.34 Proposition (Intersections of subspaces are subspaces) *If F is a field, if V is an F-vector space, and if $(\mathsf{U}_\mathrm{j})_{\mathrm{j}\in\mathrm{J}}$ is a family of subspaces, then $\cap_{\mathrm{j}\in\mathrm{J}}\mathsf{U}_\mathrm{j}$ is a subspace.*

Proof If $v\in\cap_{j\in J}\mathsf{U}_a$ and if $a\in\mathsf{F}$ then $av\in\mathsf{U}_j$ for each $j\in J$. Thus $av\in\cap_{j\in J}\mathsf{U}_j$. If $v_1,v_2\in\cap_{j\in J}\mathsf{U}_j$ then $v_1+v_2\in\mathsf{U}_j$ for each $j\in J$. Thus $v_1+v_2\in\cap_{j\in J}\mathsf{U}_j$. ■

Note that, by definition, if $(\mathsf{U}_j)_{j\in J}$ is a family of subspaces of an F-vector space V, and if $v\in\sum_{j\in J}\mathsf{U}_j$, then there exists a finite set $j_1,\dots,j_k\in J$ of indices and vectors $u_{j_l}\in\mathsf{U}_{j_l}$, $l\in\{1,\dots,k\}$, such that $v=u_{j_1}+\dots+u_{j_k}$. In taking sums of subspaces, there is an important special instance when this decomposition is unique.

4.3.35 Definition (Internal direct sum of subspaces) Let F be a field, let V be an F-vector space, and let $(\mathsf{U}_j)_{j\in J}$ be a collection of subspaces of V. The vector space V is the ***internal direct sum*** of the subspaces $(\mathsf{U}_j)_{j\in J}$, and we write $\mathsf{V} = \bigoplus_{j\in J} \mathsf{U}_j$, if, for any $v \in \mathsf{V} \setminus \{0_{\mathsf{V}}\}$, there exists unique indices $\{j_1, \dots, j_k\} \subseteq J$ and unique nonzero vectors $u_{j_l} \in \mathsf{U}_{j_l}$, $l \in \{1, \dots, k\}$, such that $v = u_{j_1} + \dots + u_{j_k}$. Each of the subspaces U_j, $j \in J$, is a ***summand*** in the internal direct sum. •

The following property of internal direct sums is useful.

4.3.36 Proposition (Representation of the zero vector in an internal direct sum of subspaces) *Let* F *be a field, let* V *be an* F*-vector space, and suppose that* V *is the internal direct sum of the subspaces* $(\mathsf{U}_j)_{j\in J}$*. If* $j_1, \dots, j_k \in J$ *are distinct and if* $u_{j_l} \in \mathsf{U}_{j_l}$, $l \in \{1, \dots, k\}$*, satisfy*

$$u_{j_1} + \dots + u_{j_k} = 0_{\mathsf{V}},$$

then $u_{j_l} = 0_{\mathsf{V}}$, $l \in \{1, \dots, k\}$.

Proof Suppose that not all of the vectors $u_{j_1}, \dots, u_{j_k}$ are zero. Without loss of generality, then, suppose that $u_{j_1} \neq 0_{\mathsf{V}}$. Then

$$u_{j_1}, \text{ and } u_{j_1} + u_{j_1} + u_{j_2} + \dots + u_{j_m} + u_{j_{m+1}}$$

are both representations of u_{j_1} as finite sums of vectors from the subspaces $(\mathsf{U}_j)_{j\in J}$. By the definition of internal direct sum it follows that $u_{j_1} = 2u_{j_1}$ and $u_{j_2} = \dots = u_{j_k} = 0_{\mathsf{V}}$. Thus $u_{j_1} = 0_{\mathsf{V}}$, which is a contradiction. ■

The following alternative characterisation of the internal direct sum is sometimes useful.

4.3.37 Proposition (Characterisation of internal direct sum for vector spaces) *Let* F *be a field, let* V *be an* F*-vector space, and let* $(\mathsf{U}_j)_{j\in J}$ *be a collection of subspaces of* V*. Then* $\mathsf{V} = \bigoplus_{j\in J} \mathsf{U}_j$ *if and only if*

(i) $\mathsf{V} = \sum_{j\in J} \mathsf{U}_j$ *and,*

(ii) for any $j_0 \in J$*, we have* $\mathsf{U}_{j_0} \cap \left(\sum_{j\in J\setminus\{j_0\}} \mathsf{U}_j\right) = \{0_{\mathsf{V}}\}$.

Proof Suppose that $\mathsf{V} = \bigoplus_{j\in J} \mathsf{U}_j$. By definition we have $\mathsf{V} = \sum_{j\in J} \mathsf{U}_j$. Let $j_0 \in J$ and suppose that $v \in \mathsf{U}_{j_0} \cap \left(\sum_{j\in J\setminus\{j_0\}} \mathsf{U}_j\right)$. Define $\mathsf{V}_{j_0} = \sum_{j\in J\setminus\{j_0\}} \mathsf{U}_j$ and note that $\mathsf{V}_{j_0} = \bigoplus_{j\in J\setminus\{j_0\}} \mathsf{U}_j$. If $v \neq 0_{\mathsf{V}}$ then there exists unique indices $j_1, \dots, j_k \in J \setminus \{j_0\}$ and unique nonzero vectors $u_{j_l} \in \mathsf{U}_{j_l}$, $l \in \{1, \dots, k\}$, such that $v = u_{j_1} + \dots + u_{j_l}$. However, since we also have $v = v$, this contradicts the fact that there exists a unique collection $j'_1, \dots, j'_{k'} \in J$ of indices and unique nonzero vectors $u_{j'_l} \in \mathsf{U}_{j'_l}$, $l' \in \{1, \dots, k'\}$, such that $v = u_{j'_1} + \dots + u_{j'_k}$. Thus we must have $v = 0_{\mathsf{V}}$.

Now suppose that (i) and (ii) hold. Let $v \in \mathsf{V} \setminus \{0_{\mathsf{V}}\}$. It is then clear from (i) that there exists indices $j_1, \dots, j_k \in J$ and nonzero vectors $u_{j_l} \in \mathsf{U}_{j_l}$, $l \in \{1, \dots, k\}$, such that $v = u_{j_1} + \dots + u_{j_k}$. Suppose that $j'_1, \dots, j'_{k'}$ and $u'_{j'_1}, \dots, u'_{j'_{k'}}$ is another collection of indices and nonzero vectors such that $v = u'_{j'_1} + \dots + u'_{j'_{k'}}$. Then

$$0_{\mathsf{V}} = u_{j_1} + \dots + u_{j_k} - (u'_{j'_1} + \dots + u'_{j'_{k'}}).$$

By Proposition 4.3.36 it follows that if $l \in \{1, \dots, k\}$ and $l' \in \{1, \dots, k'\}$ satisfy $j_l = j'_{l'}$, then $u_{j_l} = u'_{j'_{l'}}$. If for $l \in \{1, \dots, k\}$ there exists no $l' \in \{1, \dots, k'\}$ such that $j_l = j'_{l'}$, then we must have $u_{j_l} = 0_{\mathsf{V}}$. Also, if for $l' \in \{1, \dots, k'\}$ there exists no $l \in \{1, \dots, k\}$ such that $j'_{l'} = j_l$, then we must have $u'_{j'_{l'}} = 0_{\mathsf{V}}$. From this we conclude that $\mathsf{V} = \bigoplus_{j \in J} \mathsf{U}_j$. ■

The notion of internal direct sum has the following important relationship with the notion of a basis.

4.3.38 Theorem (Bases and internal direct sums for vector spaces) *Let* F *be a field, let* V *be an* F*-vector space, and let* $\mathscr{B}$ *be a basis for* V*, and define a family* $(\mathsf{U}_{\mathrm{u}})_{\mathrm{u} \in \mathscr{B}}$ *of subspaces by* $\mathsf{U}_{\mathrm{u}} = \mathrm{span}_{\mathsf{F}}(\mathrm{u})$*. Then* $\mathsf{V} = \bigoplus_{\mathrm{u} \in \mathscr{B}} \mathsf{U}_{\mathrm{u}}$*.*

Proof Let $v \in \mathsf{V}$. Since $\mathsf{V} = \mathrm{span}_{\mathsf{F}}(\mathscr{B})$, there exists $v_1, \dots, v_k \in \mathscr{B}$ and unique $c_1, \dots, c_k \in \mathsf{F}^*$ such that $v = c_1 v_1 + \dots + c_k v_k$. Therefore, $u_j = c_j v_j \in \mathsf{U}_j$ for $j \in \{1, \dots, k\}$. Thus $u_1, \dots, u_k$ are the unique nonzero elements of the subspaces $(\mathsf{U}_u)_{u \in \mathscr{B}}$ such that $v = u_1 + \dots + u_k$. ■

Let us give some examples to illustrate these manipulations involving subspaces.

4.3.39 Examples (Sums and intersections)

1. We consider a field F and the F-vector space F^3.

   (a) Let $\mathsf{U}_1 = \mathrm{span}_{\mathsf{F}}((1,0,0),(0,0,1))$ and $\mathsf{U}_2 = \mathrm{span}_{\mathsf{F}}((0,1,0))$. We claim that $\mathsf{F}^3 = \mathsf{U}_1 \oplus \mathsf{U}_2$. To see this, let $(v_1, v_2, v_3) \in \mathsf{F}^3$. Then

$$(v_1, v_2, v_3) = \underbrace{v_1(1,0,0) + v_3(0,0,1)}_{\in \mathsf{U}_1} + \underbrace{v_2(0,1,0)}_{\in \mathsf{U}_2},$$

   showing that $\mathsf{F}^3 = \mathsf{U}_1 + \mathsf{U}_2$. Moreover, if $(v_1, v_2, v_3) \in \mathsf{U}_1 \cap \mathsf{U}_2$, then $v_2 = 0_{\mathsf{F}}$ since $(v_1, v_2, v_3) \in \mathsf{U}_1$ and $v_1 = v_3 = 0_{\mathsf{F}}$ since $(v_1, v_2, v_3) \in \mathsf{F}^3$.

missing stuff

Up to this point we have considered only operations on subspaces of a given vector space. Next we consider ways of combining vector spaces that are not necessarily subspaces of a certain vector space. The reader will at this point wish to recall the notion of a general Cartesian product as given in Section 1.4.2. Much of what will be needed in these volumes relies only on finite Cartesian products, so readers not wishing to wrap their minds around the infinite case can happily consider the following constructions only for finite collections of vector spaces.

4.3.40 Definition (Direct product and direct sum of vector spaces) Let F be a field and let $(\mathsf{V}_j)_{j \in J}$ be a family of F-vector spaces.

(i) The ***direct product*** of the family $(\mathsf{V}_j)_{j \in J}$ is the F-vector space $\prod_{j \in J} \mathsf{V}_j$ with vector addition and scalar multiplication defined by

$$(f_1 + f_2)(j) = f_1(j) + f_2(j), \quad (af)(j) = a(f(j))$$

for $f, f_1, f_2 \in \prod_{j \in J} \mathsf{V}_j$ and for $a \in \mathsf{F}$.

(ii) The ***direct sum*** of the family $(\mathsf{V}_j)_{j\in J}$ is the subspace $\bigoplus_{j\in J}\mathsf{V}_j$ of $\prod_{j\in J}\mathsf{V}_j$ consisting of those elements $f\colon J\to\cup_{j\in J}\mathsf{V}_j$ for which the set $\{j\in J \mid f(j)\neq 0_{\mathsf{V}_j}\}$ is finite. Each of the vector spaces V_j, $j\in J$, is a ***summand*** in the direct sum. •

4.3.41 Notation (Finite direct products and sums) In the case when the index set J is finite, say $J=\{1,\dots,k\}$, we clearly have $\prod_{j=1}^k\mathsf{V}_j=\bigoplus_{j=1}^k\mathsf{V}_j$. We on occasion adopt the convention of writing $\mathsf{V}_1\oplus\cdots\oplus\mathsf{V}_k$ for the resulting vector space in this case. This version of the direct sum (or equivalently direct product) is the one that we will most frequently encounter. •

Let us connect the notion of a direct sum with the notion of an internal direct sum as encountered in Definition 4.3.35. This also helps to rectify the potential inconsistency of multiple uses of the symbol $\bigoplus$. The reader will want to be sure they understand infinite Cartesian products in reading this result.

4.3.42 Proposition (Internal direct sum and direct sum of vector spaces) *Let* F *be a field, let* V *be an* F*-vector space, and let* $(\mathsf{U}_\mathrm{j})_{\mathrm{j}\in\mathrm{J}}$ *be a family of subspaces of* V *such that* V *is the internal direct sum of these subspaces. Let* $\mathrm{i}_{\mathsf{U}_\mathrm{j}}\colon\mathsf{U}_\mathrm{j}\to\mathsf{V}$ *be the inclusion. Then the map from the direct sum* $\bigoplus_{\mathrm{j}\in\mathrm{J}}\mathsf{U}_\mathrm{j}$ *to* V *defined by*

$$\mathrm{f}\mapsto\sum_{\mathrm{j}\in\mathrm{J}}\mathrm{i}_{\mathsf{U}_\mathrm{j}}\mathrm{f}(\mathrm{j})$$

(noting that the sum is finite) is an isomorphism.

Proof Let us denote the map in the statement of the proposition by L. For $f,f_1,f_2\in\bigoplus_{j\in J}\mathsf{U}_j$ and for $a\in\mathsf{F}$ we have

$$\mathsf{L}(f_1+f_2)=\sum_{j\in J}(f_1+f_2)(j)=\sum_{j\in J}(f_1(j)+f_2(j))=\sum_{j\in J}f_1(j)+\sum_{j\in J}f_2(j)=\mathsf{L}(f_1)+\mathsf{L}(f_2)$$

and

$$\mathsf{L}(af)=\sum_{j\in J}(af)(j)=\sum_{j\in J}a(f(j))=a\sum_{j\in J}f(j)=a\mathsf{L}(f),$$

using the fact that all sums are finite. This proves linearity of $a\mathsf{L}$.

Next suppose that $\mathsf{L}(f)=0_{\mathsf{V}}$. By Proposition 4.3.36 it follows that $f(j)=0_{\mathsf{V}}$ for each $j\in J$. This gives injectivity of L by Exercise 4.3.23. If $v\in\mathsf{V}$, we can write $v=u_{j_1}+\cdots+u_{j_k}$ for $j_1,\dots,j_k\in J$ and for $u_{j_l}\in\mathsf{U}_{j_l}$, $l\in\{1,\dots,k\}$. If we define $f\in\bigoplus_{j\in J}\mathsf{U}_j$ by $f(j_l)=u_{j_l}$, $l\in\{1,\dots,k\}$ and $f(j)=0_{\mathsf{V}}$ for $j\notin\{j_1,\dots,j_k\}$, then $\mathsf{L}(f)=v$, showing that L is surjective. ■

4.3.43 Notation ("Internal direct sum" versus "direct sum") In the setup of the proposition, the direct sum $\bigoplus_{j\in J}\mathsf{U}_j$ is sometimes called the ***external direct sum*** of the subspaces $(\mathsf{U}_j)_{j\in J}$. The proposition says that the external direct sum is isomorphic to the internal direct sum. We shall often simply say "direct sum" rather than explicitly indicating the nature of the sum. •

Let us give an important example of a direct sum.

4.3.44 Example (The direct sum of copies of F) Let J be an arbitrary index set and let $\bigoplus_{j\in J}\mathsf{F}$ be the direct sum of "J copies" of the field F. In the case when $J = \{1,\dots,n\}$ we have $\bigoplus_{j\in J}\mathsf{F} = \mathsf{F}^n$ and in the case when $J = \mathbb{Z}_{>0}$ we have $\bigoplus_{j\in J}\mathsf{F} = \mathsf{F}_0^\infty$. Thus this example generalises two examples we have already encountered. For $j \in J$ define $\boldsymbol{e}_j\colon J \to \mathsf{F}$ by

$$\boldsymbol{e}_j(j') = \begin{cases} 1_{\mathsf{F}}, & j' = j, \\ 0_{\mathsf{F}}, & j' \neq j. \end{cases}$$

(Recall the definition of the Cartesian product to remind yourself that $\boldsymbol{e}_j \in \bigoplus_{j\in J}\mathsf{F}$.) We claim that $\{\boldsymbol{e}_j\}_{j\in J}$ is a basis for $\bigoplus_{j\in J}\mathsf{F}$. First let us show that the set is linearly independent. Let $j_1,\dots,j_k \in J$ be distinct and suppose that, for every $j' \in J$,

$$c_1\boldsymbol{e}_{j_1}(j') + \dots + c_k\boldsymbol{e}_{j_k}(j') = 0_{\mathsf{F}}$$

for some $c_1,\dots,c_k \in \mathsf{F}$. Then, taking $j' = j_l$ for $l \in \{1,\dots,k\}$ we obtain $c_l = 0_{\mathsf{F}}$. This gives linear independence. It is clear by definition of the direct sum that

$$\operatorname{span}_{\mathsf{F}}(\{\boldsymbol{e}_j\}_{j\in J}) = \bigoplus_{j\in J}\mathsf{F}.$$

We call $\{\boldsymbol{e}_j\}_{j\in J}$ the ***standard basis*** for $\bigoplus_{j\in J}\mathsf{F}$. •

4.3.45 Notation (Alternative notation for direct sums and direct products of copies of F) There will be times when it is convenient to use notation that is less transparent, but more compact, than the notation $\prod_{j\in J}\mathsf{F}$ and $\bigoplus_{j\in J}\mathsf{F}$. The notation we adopt, motivated by Examples 4.3.2–3 and 4 is

$$\prod_{j\in J}\mathsf{F} = \mathsf{F}^J, \quad \bigoplus_{j\in J}\mathsf{F} = \mathsf{F}_0^J.$$

For the direct product, this notation is in fact perfect, since, as sets, $\prod_{j\in J}\mathsf{F}$ and F^J are identical. •

The importance of the direct sum is now determined by the following theorem.

4.3.46 Theorem (Vector spaces are isomorphic to direct sums of one-dimensional subspaces) *Let F be a field, let V be an F-vector space, and let $\mathscr{B} \subseteq \mathsf{V}$ be a basis. Let $\{\mathbf{e}_{\mathrm{u}}\}_{\mathrm{u}\in\mathscr{B}}$ be the standard basis for $\bigoplus_{\mathrm{u}\in\mathscr{B}}\mathsf{F}$ and define a map $\iota_{\mathscr{B}}\colon \{\mathbf{e}_{\mathrm{u}}\}_{\mathrm{u}\in\mathscr{B}} \to \mathscr{B}$ by $\iota_{\mathscr{B}}(\mathbf{e}_{\mathrm{u}}) = \mathrm{u}$. Then there exists a unique F-isomorphism $\iota_{\mathsf{V}}\colon \bigoplus_{\mathrm{u}\in\mathscr{B}}\mathsf{F} \to \mathsf{V}$ such that the following diagram commutes:*

$$\begin{tikzcd}
\{\mathbf{e}_{\mathrm{u}}\}_{\mathrm{u}\in\mathscr{B}} \arrow[r, "\iota_{\mathscr{B}}"] \arrow[d] & \mathscr{B} \arrow[d] \\
\bigoplus_{\mathrm{u}\in\mathscr{B}}\mathsf{F} \arrow[r, "\iota_{\mathsf{V}}"'] & \mathsf{V}
\end{tikzcd}$$

where the vertical arrows represent the inclusion maps.

Proof First we define the map ι_{V}. Denote a typical element of $\bigoplus_{u\in\mathscr{B}}\mathsf{F}$ by

$$c_1\boldsymbol{e}_{u_1}+\cdots+c_k\boldsymbol{e}_{u_k}$$

for $c_1,\dots,c_k\in\mathsf{F}$ and distinct $u_1,\dots,u_k\in\mathscr{B}$. We define

$$\iota_{\mathsf{V}}(c_1\boldsymbol{e}_{u_1}+\cdots+c_k\boldsymbol{e}_{u_k})=c_1u_1+\cdots+c_ku_k.$$

It is then a simple matter to check that ι_{V} is a linear map. We also claim that it is an isomorphism. To see that it is injective suppose that

$$\iota_{\mathsf{V}}(c_1\boldsymbol{e}_{u_1}+\cdots+c_k\boldsymbol{e}_{u_k})=0_{\mathsf{V}}.$$

Then, by Proposition 4.3.36 and by the definition of ι_{V}, we have $c_1=\cdots=c_k=0_{\mathsf{F}}$. Thus the only vector mapping to zero is the zero vector, and this gives injectivity by Exercise 4.3.23. The proof of surjectivity is similarly straightforward. If $v\in\mathsf{V}$ then we can write $v=c_1u_1+\cdots+c_ku_k$ for some $c_1,\dots,c_k\in\mathsf{F}$ and $u_1,\dots,u_k\in\mathscr{B}$. Then the vector $c_1\boldsymbol{e}_{u_1}+\cdots+c_k\boldsymbol{e}_{u_k}\in\bigoplus_{u\in\mathscr{B}}\mathsf{F}$ maps to v under ι_{V}. The commutativity of the diagram in the theorem is checked directly. ■

4.3.47 Remark (Direct sums versus direct products) Note that the theorem immediately tells us that, when considering vector spaces, one can without loss of generality suppose that the vector space is a direct sum of copies of the field F. Thus direct sums are, actually, the most general form of vector space. Thinking along these lines, it becomes natural to wonder what is the value of considering direct products. First of all, Theorem 4.3.46 tells us that the direct product can be written as a direct sum, although not using the standard basis, cf. Example 4.3.28–4. The importance of the direct product will not become apparent until Section **??** when we discuss algebraic duals. •

Theorem 4.3.46 has the following corollary which tells us the relationship between the dimension of a vector space and its cardinality.

4.3.48 Corollary (The cardinality of a vector space) *If F is a field and if V is an F-vector space then*

(i) $\mathrm{card}(\mathsf{V})=\mathrm{card}(\mathsf{F})^{\dim_{\mathsf{F}}(\mathsf{V})}$ *if both* $\dim_{\mathsf{F}}(\mathsf{V})$ *and* $\mathrm{card}(\mathsf{F})$ *are finite and*

(ii) $\mathrm{card}(\mathsf{V})=\max\{\mathrm{card}(\mathsf{F}),\dim_{\mathsf{F}}(\mathsf{V})\}$ *if either* $\dim_{\mathsf{F}}(\mathsf{V})$ *or* $\mathrm{card}(\mathsf{F})$ *is infinite.*

Proof By Theorem 4.3.46, and since the dimension and cardinality of isomorphic vector spaces obviously agree (the former by Proposition 4.3.30), we can without loss of generality take the case when $\mathsf{V}=\bigoplus_{j\in J}\mathsf{F}$. We let $\{\boldsymbol{e}_j\}_{j\in J}$ be the standard basis. If J is finite then $\mathrm{card}(\mathsf{V})=\mathrm{card}(\mathsf{F})^{\mathrm{card}(J)}$ by definition of cardinal multiplication. If $\mathrm{card}(\mathsf{F})$ is finite then the result follows immediately. If $\mathrm{card}(\mathsf{F})$ is infinite then

$$\mathrm{card}(\mathsf{F})^{\mathrm{card}(J)}=\mathrm{card}(\mathsf{F})=\max\{\mathrm{card}(\mathsf{F}),\mathrm{card}(J)\}$$

by Theorem **??**. This gives the result when $\dim_{\mathsf{F}}(\mathsf{V})$ is finite.

For the case when $\mathrm{card}(J)$ is infinite, we use the following lemma.

1 Lemma *If* F *is a field and if* V *is an infinite-dimensional* F*-vector space, then* $\mathrm{card}(\mathsf{V}) = \mathrm{card}(\mathsf{F}) \cdot \dim_{\mathsf{F}}(\mathsf{V})$.

Proof As in the proof of the theorem, we suppose that $\mathsf{V} = \bigoplus_{j\in J} \mathsf{F}$. We use the fact that every vector in V is a finite linear combination of standard basis vectors. Thus

$$\mathsf{V} = \{0_{\mathsf{V}}\} \cup \Big(\cup_{k\in\mathbb{Z}_{>0}}\{c_1\boldsymbol{e}_{j_1} + \cdots + c_k\boldsymbol{e}_{j_k} \mid c_1,\dots,c_k \in \mathsf{F}^*,\ j_1,\dots,j_k \in J \text{ distinct}\}\Big). \tag{4.6}$$

Note that

$$\mathrm{card}(\{c_1\boldsymbol{e}_{j_1} + \cdots + c_k\boldsymbol{e}_{j_k} \mid c_1,\dots,c_k \in \mathsf{F}^*,\ j_1,\dots,j_k \in J \text{ distinct}\}) \\ = ((\mathrm{card}(\mathsf{F}) - 1)\,\mathrm{card}(J))^k.$$

Thus, noting that the union in (4.6) is disjoint,

$$\mathrm{card}(\mathsf{V}) = \sum_{k=0}^{\infty}((\mathrm{card}(\mathsf{F}) - 1)\,\mathrm{card}(J))^k.$$

By Theorem **??** we have

$$\mathrm{card}(\mathsf{V}) = \mathrm{card}(J)\sum_{k=0}^{\infty}(\mathrm{card}(\mathsf{F}) - 1).$$

If $\mathrm{card}(\mathsf{F})$ is finite then $\mathrm{card}(\mathsf{F}) \geq 2$ (since F contains a unit and a zero), and so, in this case, $\sum_{k=0}^{\infty}(\mathrm{card}(\mathsf{F})-1) = \mathrm{card}(\mathbb{Z}_{>0})$. If $\mathrm{card}(\mathsf{F})$ is infinite then $\sum_{k=0}^{\infty}(\mathrm{card}(\mathsf{F})-1) = \mathrm{card}(\mathsf{F})$ by Theorem **??**. In either case we have $\mathrm{card}(\mathsf{V}) = \mathrm{card}(\mathsf{F}) \cdot \mathrm{card}(J)$. ▼

We now have two cases.

1. *J is infinite and* F *is finite:* In this case we have

$$\mathrm{card}(J) \cdot \mathrm{card}(\mathsf{F}) \leq \mathrm{card}(J) \cdot \mathrm{card}(J) = \mathrm{card}(J)$$

by Theorem **??**, and we clearly have $\mathrm{card}(J) \cdot \mathrm{card}(\mathsf{F}) \geq \mathrm{card}(J)$. Thus $\mathrm{card}(J) \cdot \mathrm{card}(\mathsf{F}) = \mathrm{card}(J)$.

2. *J and* F *are both infinite:* In this case, by Theorem **??**, we have

$$\mathrm{card}(J) \cdot \mathrm{card}(\mathsf{F}) = \max\{\mathrm{card}(J), \mathrm{card}(\mathsf{F})\},$$

and the result follows. ■

We also have the following corollary to Theorem 4.3.46, along with Proposition 4.3.30, which gives an essential classification of vector spaces.

4.3.49 Corollary (Characterisation of isomorphic vector spaces) *If* F *is a field,* F*-vector spaces* V_1 *and* V_2 *are* F*-isomorphic if and only if* $\dim_{\mathsf{F}}(\mathsf{V}_1) = \dim_{\mathsf{F}}(\mathsf{V}_2)$.

Let us make Theorem 4.3.46 concrete in a simple case, just to bring things down to earth for a moment. The reader should try to draw the parallels between the relatively simple example and the more abstract proof of Theorem 4.3.46.

4.3.50 Example (Direct sum representations of finite-dimensional vector spaces) Let V be an n-dimensional vector space. By Theorem 4.3.46 we know that V is isomorphic to F^n. Moreover, the theorem explicitly indicates how an isomorphism is assigned by a basis. Thus let $\{e_1, \dots, e_n\}$ be a basis for V and let $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ be the standard basis for F^n. Then we define the map

$$\iota_{\mathscr{B}} \colon \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\} \to \{e_1, \dots, e_n\}$$

by $\iota_{\mathscr{B}}(\boldsymbol{e}_j) = e_j$, $j \in \{1, \dots, n\}$. The associated isomorphism $\iota_{\mathsf{V}} \colon \mathsf{F}^n \to \mathsf{V}$ is then given by

$$\iota_{\mathsf{V}}(v_1, \dots, v_n) = v_1 e_1 + \dots + v_n e_n.$$

The idea is simply that linear combinations of the standard basis are mapped to linear combinations of the basis for V with the coefficients preserved. •

Let us conclude our discussions in this section by understanding the relationship between direct sums and dimension. Note that, given Proposition 4.3.42, the result applies to both internal direct sums and direct sums, although it is only stated for internal direct sums.

4.3.51 Proposition (Dimension and direct sum) *Let F be a field, let V be an F-vector space, let $(\mathsf{U}_j)_{j \in J}$ be a family of F-vector spaces such that $\mathsf{V} = \bigoplus_{j \in J} \mathsf{U}_j$, and let $(\mathscr{B}_j)_{j \in J}$ be such that $\mathscr{B}_j$ is a basis for U_j. Then $\cup_{j \in J} \mathscr{B}_j$ is a basis for V. In particular,*

$$\dim_{\mathsf{F}}(\mathsf{V}) = \dim_{\mathsf{F}}(\mathsf{U}_1) + \dots + \dim_{\mathsf{F}}(\mathsf{U}_k).$$

Proof Let $v \in \mathsf{V}$. Then there exists unique $j_1, \dots, j_k \in J$ and nonzero $u_{j_l} \in \mathsf{U}_{j_l}$, $j \in \{1, \dots, k\}$, such that $v = u_{j_1} + \dots + u_{j_k}$. For each $l \in \{1, \dots, k\}$ there exists unique $c^l_1, \dots, c^l_k \in \mathsf{F}^*$ and unique $u^l_1, \dots, u^l_{k_l} \in \mathscr{B}_{j_l}$ such that

$$u_{j_l} = c^l_1 u^l_1 + \dots + c^l_{k_l} u^l_{k_l}.$$

Then we have

$$v = \sum_{l=1}^{k} \sum_{r=1}^{k_l} c^l_r u^l_r$$

as a representation of v as a finite linear combination of elements of $\cup_{j \in J} \mathscr{B}_j$ with nonzero coefficients. Moreover, this is the unique such representation since, at each step in the construction, the representations were unique. ■

4.3.6 Complements and quotients

We next consider another means of construction vector spaces from subspaces. We first address the question of when, given a subspace, there exists another subspace which gives a direct sum representation of V.

4.3.52 Definition (Complement of a subspace) If F is a field, if V is an F-vector space, and if U is a subspace of V, a ***complement*** of U in V is a subspace W of V such that $\mathsf{V} = \mathsf{U} \oplus \mathsf{W}$. •

Complements of subspaces always exist.

4.3.53 Theorem (Subspaces possess complements) *If* F *is a field, if* V *is an* F*-vector space, and if* U *is a subspace of* V*, then there exists a complement of* U*.*

Proof Let $\mathscr{B}'$ be a basis for U. By Theorem 4.3.26 there exists a basis $\mathscr{B}$ for V such that $\mathscr{B}' \subseteq \mathscr{B}$. Let $\mathscr{B}'' = \mathscr{B} \setminus \mathscr{B}'$ and define $\mathsf{W} = \operatorname{span}_{\mathsf{F}}(\mathscr{B}'')$. We claim that W is a complement of U in V. First let $v \in \mathsf{V}$. Then, since $\mathscr{B}$ is a basis for V, there exists $c'_1, \dots, c'_{k'}, c''_1, \dots, c''_{k''} \in \mathsf{F}$, $u'_1, \dots, u'_{k'} \in \mathscr{B}'$, and $u''_1, \dots, u''_{k''} \in \mathscr{B}''$ such that

$$v = \underbrace{c'_1 u'_1 + \dots + c'_{k'} u'_{k'}}_{\in \mathsf{U}} + \underbrace{c''_1 u''_1 + \dots + c''_{k''} u''_{k''}}_{\in \mathsf{W}}.$$

Thus $\mathsf{V} = \mathsf{U} + \mathsf{W}$. Next let $v \in \mathsf{U} \cap \mathsf{W}$. If $v \neq 0_{algV}$ then there exists unique $u'_1, \dots, u'_{k'} \in \mathscr{B}'$ and $u''_1, \dots, u''_{k''} \in \mathscr{B}''$ and unique $c'_1, \dots, c'_{k'}, c''_1, \dots, c''_{k''} \in \mathsf{F}$ such that

$$v = c'_1 u'_1 + \dots + c'_{k'} u'_{k'} = c''_1 u''_1 + \dots + c''_{k''} u''_{k''}.$$

This, however, contradicts the uniqueness of the representation of v as a finite linear combination of elements of $\mathscr{B}$ with nonzero coefficients. Thus $v = 0_{\mathsf{V}}$. Therefore, $\mathsf{V} = \mathsf{U} \oplus \mathsf{W}$ by Proposition 4.3.37. ■

For the same reason that a vector space possesses multiple bases, it is also the case that a strict subspace i.e., one not equal to the entire vector space, will generally possess multiple complements. Thus, while complements exist, there is not normally a natural such choice, except in the presence of additional structure (the most common such structure being an inner product, something not discussed until ***missing stuff***). However, there is a unique way in which one can associate a new vector space to a subspace in such a way that this new vector space has some properties of a complement.

4.3.54 Definition (Quotient by a subspace) Let F be a field, let V be an F-vector space, and let U be a subspace of V. The ***quotient*** of V by U is the set of equivalence classes in V under the equivalence relation

$$v_1 \sim v_2 \qquad \Longleftrightarrow \qquad v_1 - v_2 \in \mathsf{U}.$$

We denote by V/U the quotient of V by U, and we denote by $\pi_{\mathsf{V}/\mathsf{U}} \colon \mathsf{V} \to \mathsf{V}/\mathsf{U}$ the map, called the ***canonical projection***, assigning to $v \in \mathsf{V}$ its equivalence class. •

Thinking of V as an Abelian group with product defined by vector addition, the quotient V/U is simply the set of cosets of the subgroup U; see Definition 4.1.15. We shall adapt the notation for groups to denote a typical element in V/U by

$$v + \mathsf{U} = \{v + u \mid u \in \mathsf{U}\}.$$

Since V is Abelian, by Proposition 4.1.19 it follows that V/U possesses a natural Abelian group structure. It also possesses a natural vector space structure, as the following result indicates.

4.3.55 Proposition (The quotient by a subspace is a vector space) *Let* F *be a field, let* V *be an* F*-vector space, and let* U *be a subspace of* V*. The operations of vector addition and scalar multiplication in* V/U *defined by*

$$(\mathrm{v}_1 + \mathsf{U}) + (\mathrm{v}_2 + \mathsf{U}) = (\mathrm{v}_1 + \mathrm{v}_2) + \mathsf{U}, \quad \mathrm{a}(\mathrm{v} + \mathsf{U}) = (\mathrm{av}) + \mathsf{U}, \qquad \mathrm{v}, \mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V},\ \mathrm{a} \in \mathsf{F},$$

respectively, satisfy the axioms for an F*-vector space.*

Proof We define the zero vector in V/U by $0_{\mathsf{V}/\mathsf{U}} = 0_{\mathsf{V}} + \mathsf{U}$ and we define the negative of a vector $v + \mathsf{U}$ by $(-v) + \mathsf{U}$. It is then a straightforward matter to check the axioms of Definition 4.3.1, a matter which we leave to the interested reader. ■

The following "universal" property of quotients is useful.

4.3.56 Proposition (A "universal" property of quotient spaces) *Let* F *be a field, let* V *be an* F*-vector space, and let* U *be a subspace of* V*. If* W *is another* F*-vector space and if* $\mathsf{L} \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{V}; \mathsf{W})$ *has the property that* $\ker(\mathsf{L}) \subseteq \mathsf{U}$*, then there exists* $\overline{\mathsf{L}} \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{V}/\mathsf{U}; \mathsf{W})$ *such that the diagram*

$$\begin{tikzcd}
\mathsf{V} \arrow[r, "\mathsf{L}"] \arrow[d, "\pi_{\mathsf{V}/\mathsf{U}}"'] & \mathsf{W} \\
\mathsf{V}/\mathsf{U} \arrow[ur, "\overline{\mathsf{L}}"'] &
\end{tikzcd}$$

commutes. Moreover, if $\overline{\mathsf{L}}' \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{V}/\mathsf{U}; \mathsf{W})$ *is such that the preceding diagram commutes, then* $\overline{\mathsf{L}}' = \overline{\mathsf{L}}$*.*

Proof We define $\overline{\mathsf{L}}(v + \mathsf{U}) = \mathsf{L}(v)$. This map is well-defined since, if $v' + \mathsf{U} = v + \mathsf{U}$ then $v' = v + u$ for $u \in \mathsf{U}$, whence

$$\overline{\mathsf{L}}(v' + \mathsf{U}) = \mathsf{L}(v') = \mathsf{L}(v + u) = \mathsf{L}(v) = \overline{\mathsf{L}}(v + \mathsf{U}).$$

One verifies directly that

$$\overline{\mathsf{L}}((v_1 + \mathsf{U}) + (v_2 + \mathsf{U})) = \overline{\mathsf{L}}(v_1 + \mathsf{U}) + \overline{\mathsf{L}}(v_2 + \mathsf{U}), \qquad \overline{\mathsf{L}}(a(v + \mathsf{U})) = a\overline{\mathsf{L}}(v + \mathsf{U}),$$

giving linearity of $\overline{\mathsf{L}}$. For the final assertion of the proposition, the commuting of the diagram exactly says that $\overline{\mathsf{L}}'(v + \mathsf{U}) = \mathsf{L}(v)$, as desired. ■

Next we consider the relationship between complements and quotient spaces.

4.3.57 Theorem (Relationship between complements and quotients) *Let* F *be a field, let* V *be an* F*-vector space, and let* U *be a subspace of* V *with a complement* W*. Then the map* $\iota_{\mathsf{U},\mathsf{W}} \colon \mathsf{W} \to \mathsf{V}/\mathsf{U}$ *defined by*

$$\iota_{\mathsf{U},\mathsf{W}}(\mathrm{w}) = \mathrm{w} + \mathsf{U}$$

is an isomorphism. In particular, $\dim_{\mathsf{F}}(\mathsf{W}) = \dim_{\mathsf{F}}(\mathsf{V}/\mathsf{U})$ *for any complement* W *of* U *in* V*.*

Proof The map $\iota_{\mathsf{U},\mathsf{W}}$ is readily checked to be linear, and we leave this verification to the reader. Suppose that $w + \mathsf{U} = 0_{\mathsf{V}} + \mathsf{U}$ for $w \in \mathsf{W}$. This implies that $w \in \mathsf{U}$, which gives $w = 0_{\mathsf{V}}$ by Proposition 4.3.37; thus $\iota_{\mathsf{U},\mathsf{W}}$ is injective by Exercise 4.3.23. Now let $v + \mathsf{U} \in \mathsf{V}/\mathsf{U}$. Since $\mathsf{V} = \mathsf{U} \oplus \mathsf{W}$ we can write $v = u + w$ for $u \in \mathsf{U}$ and $w \in \mathsf{W}$. Since $v - w \in \mathsf{U}$ we have $v + \mathsf{U} = w + \mathsf{U}$. Thus $\iota_{\mathsf{U},\mathsf{W}}$ is also surjective.

The final assertion follows from Propositions 4.3.30 and 4.3.51. ■

The preceding result gives the dimension of the quotient, and the next result reinforces this by giving an explicit basis for the quotient.

4.3.58 Proposition (Basis for quotient) *Let* F *be a field, let* V *be an* F*-vector space, and let* U *be a subspace of* V*. If* $\mathscr{B}$ *is a basis for* V *with the property that there exists a subset* $\mathscr{B}' \subseteq \mathscr{B}$ *with the property that* $\mathscr{B}'$ *is a basis for* U*, then*

$$\{\mathrm{v} + \mathsf{U} \mid \mathrm{v} \in \mathscr{B} \setminus \mathscr{B}'\}$$

is a basis for V/U*.*

Proof Let $\mathscr{B}''$ be such that $\mathscr{B} = \mathscr{B}' \cup \mathscr{B}''$ and $\mathscr{B}' \cup \mathscr{B}'' = \emptyset$. If $v \in \mathsf{V}$ then we can write

$$v = c_1u_1 + \cdots + c_ku_k + d_1v_1 + \cdots + d_lv_l$$

for $c_1, \dots, c_k, d_1, \dots, d_l \in \mathsf{F}$, for $u_1, \dots, u_k \in \mathscr{B}'$, and for $v_1, \dots, v_l \in \mathscr{B}''$. Then

$$\begin{aligned} v + \mathsf{U} &= (c_1u_1 + \cdots + c_ku_k + d_1v_1 + \cdots + d_lv_l) + \mathsf{U} \\ &= (d_1v_1 + \cdots + d_lv_l) + \mathsf{U} = (d_1v_1 + \mathsf{U}) + \cdots + (d_lv_l + \mathsf{U}), \end{aligned}$$

showing that $\{v + \mathsf{U} \mid v \in \mathscr{B}''\}$ generates V/U. To show linear independence, suppose that

$$(d_1v_1 + \mathsf{U}) + \cdots + (d_lv_l + \mathsf{U}) = 0_{\mathsf{V}} + \mathsf{U}$$

for $v_1, \dots, v_l \in \mathscr{B}''$ and $d_1, \dots, d_l \in \mathsf{F}$. Then $d_1v_1 + \cdots + d_lv_l \in \mathsf{U}$, and so $d_1v_1 + \cdots + d_lv_l = 0_{\mathsf{V}}$ by Proposition 4.3.37. Since $\mathscr{B}''$ is linearly independent by Proposition 4.3.19(iii), it follows that $d_1 = \cdots = d_l = 0_{\mathsf{F}}$, and so $\{v + \mathsf{U} \mid v \in \mathscr{B}''\}$ is linearly independent. ■

The preceding theorem motivates the following definition.

4.3.59 Definition (Codimension of a subspace) Let F be a field, let V be an F-vector space, and let U be a subspace of V. The ***codimension*** of U, denoted by $\operatorname{codim}_{\mathsf{F}}(\mathsf{U})$, is $\dim_{\mathsf{F}}(\mathsf{V}/\mathsf{U})$. •

Combining Proposition 4.3.51 and Theorem 4.3.57 immediately gives the following result.

4.3.60 Corollary (Dimension and codimension of a subspace) *If* F *is a field, if* V *is an* F*-vector space, and if* U *is a subspace of* V*, then* $\dim_{\mathsf{F}}(\mathsf{V}) = \dim_{\mathsf{F}}(\mathsf{U}) + \operatorname{codim}_{\mathsf{F}}(\mathsf{U})$*.*

4.3.7 Complexification of $\mathbb{R}$-vector spaces

It will often be useful to regard a vector space defined over $\mathbb{R}$ as being defined over $\mathbb{C}$. This is fairly straightforward to do.

4.3.61 Definition (Complexification of a $\mathbb{R}$-vector space) If V is a $\mathbb{R}$-vector space, the ***complexification*** of V is the $\mathbb{C}$-vector space $\mathsf{V}_{\mathbb{C}}$ defined by

(i) $\mathsf{V}_{\mathbb{C}} = \mathsf{V} \times \mathsf{V}$,

and with the operations of vector addition and scalar multiplication defined by

(ii) $(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$, $u_1, u_2, v_1, v_2 \in \mathsf{V}$, and

(iii) $(a + \mathrm{i}b)(u, v) = (au - bv, av + bu)$ for $a, b \in \mathbb{R}$ and $u, v \in \mathsf{V}$. •

We recall from Example 4.3.2–5 that any $\mathbb{C}$-vector space is also a $\mathbb{R}$-vector space by simply restricting scalar multiplication to $\mathbb{R}$. It will be convenient to regard V as a subspace of the $\mathbb{R}$-vector space $\mathsf{V}_{\mathbb{C}}$. There are many ways one might do this. For example, we can identify V with the either of the two subspaces

$$\{(u, v) \in \mathsf{V}_{\mathbb{C}} \mid v = 0_{\mathsf{V}}\}, \quad \{(u, v) \in \mathsf{V}_{\mathbb{C}} \mid u = 0_{\mathsf{V}}\},$$

and there are many other possible choices. However, the subspace on the left is the most natural one for reasons that will be clear shortly. We thus define the monomorphism $\iota_{\mathsf{V}} \colon \mathsf{V} \to \mathsf{V}_{\mathbb{C}}$ of $\mathbb{R}$-vector spaces by $\iota(v) = (v, 0_{\mathsf{V}})$, and we note that $\operatorname{image}(\iota_{\mathsf{V}})$ is a subspace of $\mathsf{V}_{\mathbb{C}}$ that is isomorphic to V.

The following result records that $\mathsf{V}_{\mathbb{C}}$ has the desired properties.

4.3.62 Proposition (Properties of complexification) *If V is a $\mathbb{R}$-vector space then the complexification $\mathsf{V}_{\mathbb{C}}$ has the following properties:*

(i) $\mathsf{V}_{\mathbb{C}}$ is a $\mathbb{C}$-vector space and $\dim_{\mathbb{C}}(\mathsf{V}_{\mathbb{C}}) = \dim_{\mathbb{R}}(\mathsf{V})$;

(ii) $\mathsf{V}_{\mathbb{C}}$ is a $\mathbb{R}$-vector space and $\dim_{\mathbb{R}}(\mathsf{V}_{\mathbb{C}}) = 2\dim_{\mathbb{R}}(\mathsf{V})$;

(iii) every element of $\mathsf{V}_{\mathbb{C}}$ can be uniquely expressed as $\iota_{\mathsf{V}}(\mathrm{u}) + \mathrm{i}\,\iota_{\mathsf{V}}(\mathrm{v})$ for some $\mathrm{u}, \mathrm{v} \in \mathsf{V}$.

Proof (i) The verification of the axioms for $\mathsf{V}_{\mathbb{C}}$ to be a $\mathbb{C}$-vector space is straightforward and relatively unilluminating, so we leave the reader to fill in the details. Let us verify that $\dim_{\mathbb{C}}(\mathsf{V}) = \dim_{\mathbb{R}}(\mathsf{V})$. Let $\mathscr{B}$ be a basis for V and define

$$\mathscr{B}_{\mathbb{C}} = \{(u, 0_{\mathsf{V}}) \mid u \in \mathscr{B}\}.$$

We claim that $\mathscr{B}_{\mathbb{C}}$ is a basis for $\mathsf{V}_{\mathbb{C}}$ as a $\mathbb{C}$-vector space. To show linear independence of $\mathscr{B}_{\mathbb{C}}$, suppose that

$$(a_1 + \mathrm{i}b_1)(u_1, 0_{\mathsf{V}}) + \cdots + (a_k + \mathrm{i}b_k)(u_k, 0_{\mathsf{V}}) = (0_{\mathsf{V}}, 0_{\mathsf{V}})$$

for $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{R}$. Using the definition of scalar multiplication this implies that

$$(a_1u_1, b_1u_1) + \cdots + (a_ku_k, b_ku_k) = (0_{\mathsf{V}}, 0_{\mathsf{V}}).$$

Linear independence of $\mathscr{B}$ then implies that $a_j = b_j = 0$ for $j \in \{1, \dots, k\}$, so giving linear independence of $\mathscr{B}_{\mathbb{C}}$. Now let $(u, v) \in \mathsf{V}_{\mathbb{C}}$. There then exists $u_1, \dots, u_k \in \mathscr{B}$ and $a_1, \dots, a_k, b_1, \dots, b_k \in \mathbb{R}$ such that

$$u = a_1u_1 + \cdots + a_ku_k, \quad v = b_1u_1 + \cdots + b_ku_k.$$

We then have

$$(u, v) = (a_1u_1 + \cdots + a_ku_k, b_1u_1 + \cdots + b_ku_k) = (a_1u_1, b_1u_1) + \cdots + (a_ku_k, b_ku_k).$$

Using the rules for scalar multiplication in $\mathsf{V}_{\mathbb{C}}$ this gives

$$(u,v) = (a_1 + \mathrm{i}b_1)(u_1, 0_{\mathsf{V}}) + \cdots + (a_k + \mathrm{i}b_k)(u_k, 0_{\mathsf{V}}).$$

Thus $\mathscr{B}_{\mathbb{C}}$ spans $\mathsf{V}_{\mathbb{C}}$, and so is a basis for $\mathsf{V}_{\mathbb{C}}$.

(ii) That $\mathsf{V}_{\mathbb{C}}$ is a $\mathbb{R}$-vector space follows from Example 4.3.2–5. Note that scalar multiplication in the $\mathbb{R}$-vector space $\mathsf{V}_{\mathbb{C}}$, i.e., restriction of $\mathbb{C}$ scalar multiplication to $\mathbb{R}$, is defined by $a(u,v) = (au, av)$. Thus $\mathsf{V}_{\mathbb{C}}$ as a $\mathbb{R}$-vector space is none other than $\mathsf{V} \oplus \mathsf{V}$. That $\dim_{\mathbb{R}}(\mathsf{V}_{\mathbb{C}}) = 2\dim_{\mathbb{R}}(\mathsf{V})$ then follows from Proposition 4.3.51.

(iii) Using the definition of $\mathbb{C}$ scalar multiplication we have

$$\mathrm{i}\iota_{\mathsf{V}}(v) = \mathrm{i}(v, 0_{\mathsf{V}}) = (0_{\mathsf{V}}, v).$$

Thus we clearly have

$$(u,v) = \iota_{\mathsf{V}}(u) + \mathrm{i}\,\iota_{\mathsf{V}}(v),$$

giving the existence of the stated representation. Now, if

$$\iota_{\mathsf{V}}(u_1) + \mathrm{i}\,\iota_{\mathsf{V}}(v_1) = \iota_{\mathsf{V}}(u_2) + \mathrm{i}\,\iota_{\mathsf{V}}(v_2),$$

then $(u_1, v_1) = (u_2, v_2)$, and so $u_1 = u_2$ and $v_1 = v_2$, giving uniqueness of the representation. ■

The final assertion in the proposition says that we can think of $(u,v) \in \mathsf{V}_{\mathbb{C}}$ as $(u, 0_{\mathsf{V}}) + \mathrm{i}(v, 0_{\mathsf{V}})$. With this as motivation, we shall use the notation $(u,v) = u + \mathrm{i}v$ when it is convenient. This then leads to the following definitions which adapt those for complex numbers to the complexification of a $\mathbb{R}$-vector space.

4.3.63 Definition (Real part, imaginary part, complex conjugation) Let V be a $\mathbb{R}$-vector space with $\mathsf{V}_{\mathbb{C}}$ its complexification.

(i) The ***real part*** of $(u,v) \in \mathsf{V}_{\mathbb{C}}$ is $\mathrm{Re}(u,v) = u$.

(ii) The ***imaginary part*** of $(u,v) \in \mathsf{V}_{\mathbb{C}}$ is $\mathrm{Im}(u,v) = v$.

(iii) The representation $u + \mathrm{i}v$ of $(u,v) \in \mathsf{V}_{\mathbb{C}}$ is the ***canonical representation***.

(iv) ***Complex conjugation*** is the map $\sigma_{\mathsf{V}}\colon \mathsf{V}_{\mathbb{C}} \to \mathsf{V}_{\mathbb{C}}$ defined by $\sigma_{\mathsf{V}}(u,v) = (u, -v)$. •

Using the canonical representation of elements in the complexification, $\mathbb{C}$-scalar multiplication in $\mathsf{V}_{\mathbb{C}}$ can be thought of as applying the usual rules for $\mathbb{C}$ multiplication to the expression $(a + \mathrm{i}b)(u + \mathrm{i}v)$:

$$(a + \mathrm{i}b)(u + \mathrm{i}v) = (au - bv) + \mathrm{i}(bu + av).$$

This is a helpful mnemonic for remembering the scalar multiplication rule for $\mathsf{V}_{\mathbb{C}}$.

It is easy to show that $\sigma_{\mathsf{V}} \in \mathrm{End}_{\mathbb{R}}(\mathsf{V}\mathbb{C})$, but that $\sigma_{\mathsf{V}} \notin \mathrm{End}_{\mathbb{C}}(\mathsf{V}_{\mathbb{C}})$ (see Exercise 4.3.25). Moreover, complex conjugation has the following easily verified properties.

The following example should be thought of, at least in the finite-dimensional case, as the typical one.

4.3.64 Example ($\mathbb{R}^n_{\mathbb{C}} = \mathbb{C}^n$) We take the $\mathbb{R}$-vector space $\mathbb{R}^n$ and consider its complexification $\mathbb{R}^n_{\mathbb{C}}$. The main point to be made here is the following lemma.

1 **Lemma** *The map* $(\mathrm{x}_1, \dots, \mathrm{x}_\mathrm{n}) + \mathrm{i}(\mathrm{y}_1, \dots, \mathrm{y}_\mathrm{n}) \mapsto (\mathrm{x}_1 + \mathrm{i}\mathrm{y}_1, \dots, \mathrm{x}_\mathrm{n} + \mathrm{i}\mathrm{y}_\mathrm{n})$ *is a* $\mathbb{C}$*-isomorphism of* $\mathbb{R}^{\mathrm{n}}_{\mathbb{C}}$ *with* $\mathbb{C}^{\mathrm{n}}$*.*

Proof This follows by the definition of vector addition and $\mathbb{C}$-scalar multiplication in $\mathbb{R}^n_{\mathbb{C}}$. ▼

Let us look at some of the constructions associated with complexification in order to better understand them. First note that $\mathbb{R}^n_{\mathbb{C}}$ has the structure of both a $\mathbb{R}$- and $\mathbb{C}$-vector space. One can check that a basis for $\mathbb{R}^n_{\mathbb{C}}$ as a $\mathbb{R}$-vector space is given by the set

$$\{\boldsymbol{e}_1 + \mathrm{i}\boldsymbol{0}, \dots, \boldsymbol{e}_n + \mathrm{i}\boldsymbol{0}, \boldsymbol{0} + \mathrm{i}\boldsymbol{e}_1, \dots, \boldsymbol{0} + \mathrm{i}\boldsymbol{e}_n\},$$

and a basis for $\mathbb{R}^n_{\mathbb{C}}$ as a $\mathbb{C}$-vector space is given by the set

$$\{\boldsymbol{e}_1 + \mathrm{i}\boldsymbol{0}, \dots, \boldsymbol{e}_n + \mathrm{i}\boldsymbol{0}\},$$

where $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is the standard basis for $\mathbb{R}^n$. It is also clear that

$$\mathrm{Re}(\boldsymbol{x} + \mathrm{i}\boldsymbol{y}) = \boldsymbol{x}, \quad \mathrm{Im}(\boldsymbol{x} + \mathrm{i}\boldsymbol{y}) = \boldsymbol{y}, \quad \sigma_{\mathbb{R}^n}(\boldsymbol{x} + \mathrm{i}\boldsymbol{y}) = \boldsymbol{x} - \mathrm{i}\boldsymbol{y}.$$

The idea in this example is, essentially, that one can regard the complexification of $\mathbb{R}^n$ as the vector space obtained by "replacing" the real entries in a vector with complex entries. •

4.3.8 Extending the scalars for a vector space

In Section 4.3.7 we saw how one can naturally regard a $\mathbb{R}$-vector space as a $\mathbb{C}$-vector space. In this section we generalise this idea to general field extensions, as it will be useful in studying endomorphisms of finite-dimensional vector spaces in Section **??**. This development relies on the tensor product which itself is a part of multilinear algebra. Thus a reader will need to make a diversion ahead to Section **??** in order to understand the material in this section.

While we have not yet discussed field extensions (we do so formally and in detail in Section **??**), the notion is a simple one. A field K that contains a field F as a subfield is an ***extension*** of F. As we will show in Proposition **??**, and is easily seen in any case, K is an F-vector space. We shall make essential use of this fact in this section. Indeed, the key idea in complexification comes from understanding the $\mathbb{R}$-vector space structure of $\mathbb{C}$. Here we generalise this idea.

We may now define the extension of an F-vector space to an extension K of F. This definition will seem odd at first glance, relying as it does on the tensor product. It is only after we explore it a little that it will (hopefully) seem "correct."

4.3.65 **Definition (Extension of scalars for a vector space)** Let F be a field, let K be an extension of F, and let V be an F-vector space. The ***extension*** of V to K is

$$\mathsf{V}_{\mathsf{K}} = \mathsf{K} \otimes \mathsf{V}.$$ •

At this point, we certainly understand all the symbols in the definition. However, it is not so clear what V_{K} really is. To begin to understand it, let us first show that it has the structure of a vector space over K; it is this structure that is of most interest to us.

4.3.66 Proposition (V_K is an K-vector space) *Let* K *be an extension of a field* F *and let* V *be an* F*-vector space. Using vector addition and scalar multiplication defined by vector addition in* $\mathsf{K} \otimes \mathsf{V}$ *(as an* F*-vector space) and* $\mathrm{b}(\mathrm{a} \otimes \mathrm{v}) = (\mathrm{ab}) \otimes \mathrm{v}$, $\mathrm{a}, \mathrm{b} \in \mathsf{K}$, $\mathrm{v} \in \mathsf{V}$, *respectively,* $\mathsf{K} \otimes \mathsf{V}$ *is a vector space over* K*.*

Proof First let us show that the definition of scalar multiplication in K is well-defined. We note that for $b \in \mathsf{K}$ the map $\phi_b \colon \mathsf{K} \times \mathsf{V} \to \mathsf{K} \otimes \mathsf{V}$ defined by $\phi_b(a, v) = (ba) \otimes v$ is bilinear. Thus there exists a unique linear map $\mathsf{L}_{\phi_b} \colon \mathsf{K} \otimes \mathsf{V} \to \mathsf{K} \otimes \mathsf{V}$ satisfying $\mathsf{L}_{\phi_b}(a \otimes v) = (ba) \otimes v$. Now, if

$$a_1 \otimes v_1 + \cdots + a_k \otimes v_k$$

is an arbitrary element of $\mathsf{K} \otimes \mathsf{V}$, it follows that

$$\mathsf{L}_{\phi_b}(a_1 \otimes v_1 + \cdots + a_k \otimes v_k) = (ba_1) \otimes v_1 + \cdots + (ba_k) \otimes v_k$$

since L_{ϕ_b} is linear. Thus scalar multiplication is well-defined on all of $\mathsf{K} \otimes \mathsf{V}$. To show that vector addition and scalar multiplication satisfy the usual axioms for a vector space is now straightforward, and we leave the details of this to the reader. ■

Let us show that this complicated notion of scalar extension agrees with complexification.

4.3.67 Example ($\mathsf{V}_\mathbb{C} = \mathbb{C} \otimes \mathsf{V}$) We let V be a $\mathbb{R}$-vector space with complexification $\mathsf{V}_\mathbb{C}$. Let us show that "$\mathsf{V}_\mathbb{C} = \mathsf{V}_\mathbb{C}$;" i.e., that complexification as in Section 4.3.7 agrees with extension of scalars as in Definition 4.3.65. To see this we define an isomorphism $\iota_\mathbb{C}$ from $\mathsf{V}_\mathbb{C}$ (the complexification as in Section 4.3.7) to $\mathbb{C} \otimes \mathsf{V}$ by

$$\iota_\mathbb{C}(u, v) = 1 \otimes u + \mathrm{i} \otimes v.$$

Let us show that this is an isomorphism of $\mathbb{C}$-vector spaces. First we note that

$$\begin{aligned}\iota_\mathbb{C}((u_1, v_1) + (u_2, v_2)) &= \iota_\mathbb{C}(u_1 + u_2, v_1 + v_2) = 1 \otimes (u_1 + u_2) + \mathrm{i}(v_1 + v_2)\\ &= (1 \otimes u_1 + \mathrm{i}v_1) + (1 \otimes u_2 + \mathrm{i} \otimes v_2) = \iota_\mathbb{C}(u_1, v_1) + \iota_\mathbb{C}(u_2, v_2)\end{aligned}$$

and

$$\begin{aligned}\iota_\mathbb{C}((a + \mathrm{i}b)(u, v)) &= \iota_\mathbb{C}(au - bv, av + bu) = 1 \otimes (au - bv) + \mathrm{i}(av + bu)\\ &= 1 \otimes (au) + 1 \otimes (-bv) + \mathrm{i} \otimes (av) + \mathrm{i} \otimes (bu)\\ &= a \otimes u + (-b) \otimes v + (\mathrm{i}a) \otimes v + (\mathrm{i}b) \otimes u\\ &= a(1 \otimes u + \mathrm{i} \otimes v) + \mathrm{i}b(1 \otimes u + \mathrm{i} \otimes v)\\ &= (a + \mathrm{i}b)(1 \otimes u + \mathrm{i} \otimes v) = (a + \mathrm{i}b)\iota_\mathbb{C}(u, v),\end{aligned}$$

so showing that $\iota_\mathbb{C}$ is a $\mathbb{C}$-linear. To show that $\iota_\mathbb{C}$ is injective, suppose that $\iota_\mathbb{C}(u, v) = 0_{\mathbb{C} \otimes \mathsf{V}}$. Thus

$$1 \otimes u + \mathrm{i} \otimes v = 1 \otimes 0_\mathsf{V} + \mathrm{i} \otimes 0_\mathsf{V},$$

and so $u = v = 0_\mathsf{V}$. Thus $\iota_\mathbb{C}$ is injective by Exercise 4.3.23. To show that $\iota_\mathbb{C}$ is surjective, it suffices (why?) to show that $(a + \mathrm{i}b) \otimes v \in \mathrm{image}(\iota_\mathbb{C})$ for each $a, b \in \mathbb{R}$ and $v \in \mathsf{V}$. This follows since

$$\iota_\mathbb{C}(av, bv) = 1 \otimes (av) + \mathrm{i} \otimes (bv) = a \otimes v + (\mathrm{i}b) \otimes v = (a + \mathrm{i}b) \otimes v.$$

Note that $1 \otimes u + \mathrm{i} \otimes v$ is the corresponding decomposition of $(u, v) \in \mathsf{V}_{\mathbb{C}}$ into its real and imaginary parts. If one keeps this in mind, and uses the usual rules for manipulating tensor products, it is easy to see why $\mathbb{C} \otimes \mathsf{V}$ is, indeed, the complexification of V. •

4.3.9 Notes

Exercises

4.3.1 Verify the vector space axioms for Example 4.3.2–1.

4.3.2 Verify the vector space axioms for Example 4.3.2–2.

4.3.3 Verify the vector space axioms for Example 4.3.2–3.

4.3.4 Verify the vector space axioms for Example 4.3.2–4.

4.3.5 Verify the vector space axioms for Example 4.3.2–5.

4.3.6 Verify the vector space axioms for Example 4.3.2–6.

4.3.7 Verify the vector space axioms for Example 4.3.2–7.

4.3.8 Verify the vector space axioms for Example 4.3.2–8.

4.3.9 Verify the vector space axioms for Example 4.3.2–9.

4.3.10 Let $I \subseteq \mathbb{R}$, let $r \in \mathbb{Z}_{>0}$, and denote by $\mathsf{C}^r(I; \mathbb{R})$ the set of $\mathbb{R}$-valued functions on I that are r-times continuously differentiable. Define vector addition and scalar multiplication in such a way that $\mathsf{C}^r(I; \mathbb{R})$ is a $\mathbb{R}$-vector space.

4.3.11 Prove Proposition 4.3.6.

4.3.12 Verify the claim of Example 4.3.7–1.

4.3.13 Verify the claim of Example 4.3.7–2.

4.3.14 Verify the claim of Example 4.3.7–3.

4.3.15 Verify the claim of Example 4.3.7–4.

4.3.16 Prove Proposition 4.3.9.

4.3.17 Do the following.

(a) Give an example of a vector space V and two subspaces U_1 and U_2 of V such that $\mathsf{U}_1 \cup \mathsf{U}_2$ is not a subspace.

(b) If V is an F-vector space and if $\mathsf{U}_1, \dots, \mathsf{U}_k$ are subspaces of V, show that $\cup_{j=1}^{k} \mathsf{U}_j$ is a subspace if and only if there exists $j_0 \in \{1, \dots, k\}$ such that $\mathsf{U}_j \subseteq \mathsf{U}_{j_0}$ for $j \in \{1, \dots, k\}$.

(c) If V is an F-vector space and if $(\mathsf{U}_j)_{j \in J}$ is an arbitrary family of subspaces, give conditions, analogous to those of part (b), that ensure that $\cup_{j \in J} \mathsf{U}_j$ is a subspace.

4.3.18 Prove Theorem 4.3.26 in the case when $\dim_{\mathsf{F}}(\mathsf{V}) < \infty$.

4.3.19 Let F be a field, let $\mathsf{V}$and W be F-vector spaces, let $\mathscr{B} \subseteq \mathsf{V}$ be a basis, let $\phi \colon \mathscr{B} \to \mathsf{W}$ be a map, and let $\mathsf{L}_\phi \in \mathrm{Hom}_{\mathsf{F}}(\mathsf{V}; \mathsf{W})$ be the unique linear map determined as in Theorem 4.3.24.

(a) Show that L_ϕ is injective if and only if the family $(\phi(v))_{v \in \mathscr{B}}$ is linearly independent.

(b) Show that L_ϕ is surjective if and only if $\operatorname{span}_{\mathsf{F}}(\phi(\mathscr{B})) = \mathsf{W}$.

4.3.20 Let F be a field and let V be an F-vector space. If U is a subspace of V and if $v_1, v_2 \in \mathsf{V}$, show that the affine subspaces

$$\{v_1 + u \mid u \in \mathsf{U}\}, \quad \{v_2 + u \mid u \in \mathsf{U}\}$$

agree if and only if $v_1 - v_2 \in \mathsf{U}$.

4.3.21 Construct explicit isomorphisms between the following pairs of F-vector spaces:

(a) F^{k+1} and $\mathsf{F}_k[\xi]$;

(b) F_0^∞ and $\mathsf{F}[\xi]$.

4.3.22 Construct an explicit $\mathbb{R}$-isomorphism between $\mathbb{R}^\infty$ and the set $\mathbb{R}[[\xi]]$ of $\mathbb{R}$-formal power series.

4.3.23 Let F be a field, let U and V be F-vector spaces, and let $\mathsf{L} \in \operatorname{Hom}_{\mathsf{F}}(\mathsf{U}; \mathsf{V})$. Show that L is injective if and only if $\ker(\mathsf{L}) = \{0_{\mathsf{U}}\}$.

4.3.24 Let F be a field and let V be an F-vector space with U a strict subspace of V.

(a) Show that, if $\dim_{\mathsf{F}}(\mathsf{V}) < \infty$, then $\dim_{\mathsf{F}}(\mathsf{U}) < \dim_{\mathsf{F}}(\mathsf{V})$.

(b) Give examples of F, V, and U as above such that $\dim_{\mathsf{F}}(\mathsf{U}) = \dim_{\mathsf{F}}(\mathsf{V})$.

4.3.25 Let V be a $\mathbb{R}$-vector space with $\mathsf{V}_{\mathbb{C}}$ its complexification. Show that the complex conjugation σ_{V} is a $\mathbb{R}$-linear map of $\mathsf{V}_{\mathbb{C}}$, but not a $\mathbb{C}$-linear map.

Chapter 5

Measure theory and integration

The theory of measure and integration we present in this chapter represents one of the most important achievements of mathematics in the twentieth century. To a newcomer to the subject or to someone coming at the material from an "applied" perspective, it can be difficult to understand *why* abstract integration provides anything of value. This is the more so if one comes equipped with the knowledge of Riemann integration as we have developed in Sections 3.4 and **??**. This theory of integration appears to be entirely satisfactory. There are certainly functions that are easily described, but not Riemann integrable (see Example 3.4.10). However, these functions typically fall into the class of functions that one will not encounter in practice, so it is not clear that they represent a serious obstacle to the viability of Riemann integration. Indeed, if one's objective is only to compute integrals, then the Riemann integral is all that is needed. The multiple volumes of tables of integrals, many of them several hundred pages in length, are all compiled using good ol' Riemann integration. *But this is not the problem that is being addressed by modern integration theory!* The theory of measure and integration we present in this chapter is intended to provide a theory whereby *spaces* of integrable functions have satisfactory properties. This confusion concerning the objectives of modern integration theory is widespread. For example, an often encountered statement is that of Richard W. Hamming (1915–1998):

> Does anyone believe that the difference between the Lebesgue and Riemann integrals can have physical significance, and that whether say, an airplane would or would not fly could depend on this difference? If such were claimed, I should not care to fly in that plane.

We are uncertain what Hamming was actually saying when he made this statement. However, it is certainly the case that this statement gets pulled out by many folks as justification for the statement that the modern theory of integration is simply not worth learning. Our view on this is that it may well be the case that this is true. If all you want to be able to do is integrate functions, then there is no need to learn the modern theory of integration. However, if you find yourself talking about spaces of integrable functions (as we shall do constantly in Volume **??*missing stuff*** in our discussion of signal theory), then you will find yourself needing a theory of integration that is better that Riemann integration.

With the above as backdrop, in Section 5.1 we discuss in detail some of the limitations of the Riemann integral. After doing this we launch into a treatment of measure theory and integration. While there is no question that the special case of Lebesgue measure and integration is of paramount importance for us, we take the approach that measure theory and integration is actually easier to understand starting from a general point of view. Thus we start with general measure theory and the corresponding general integration theory. We then specialise to Lebesgue measure and integration.

Do I need to read this chapter? The reader ought to be able to decide based on the discussion above whether they want to read this chapter. If they elect to bypass it, then they will be directed back to it at appropriate points in the sequel.

That being said, it is worth attempting to disavow a common perception about the use of measure theory and integration. There appears to be a common feeling that the theory is difficult, weird, and overly abstract. Part of this may stem from the fact that many already have a comfort level with integration via the Riemann integral, and so do not feel compelled to relearn integration theory. But the fact is that measure theory is no more difficult to learn than anything else about real analysis. •

Contents

5.1 Some motivation for abstract measure theory and integration . . . 347
5.1.1 The Jordan measure and its limitations . . . 347
5.1.2 Some limitations of the Riemann integral . . . 356
5.1.3 An heuristic introduction to the Lebesgue integral . . . 359
5.1.4 Notes . . . 361
5.2 Measurable sets . . . 362
5.2.1 Algebras and σ-algebras . . . 362
5.2.2 Algebras and σ-algebras generated by families of subsets . . . 364
5.2.3 Products of measurable spaces . . . 369
5.3 Measures . . . 374
5.3.1 Functions on families of subsets . . . 374
5.3.2 Outer measures, measures, and their relationship . . . 379
5.3.3 Complete measures and completions of measures . . . 387
5.3.4 Outer and inner measures associated to a measure . . . 389
5.3.5 Probability measures . . . 396
5.3.6 Product measures . . . 398
5.3.7 Signed measures . . . 404
5.3.8 Complex measures . . . 411
5.3.9 Vector measures . . . 414
5.3.10 Spaces of positive, signed, complex, and vector measures . . . 416
5.3.11 Notes . . . 418
5.4 Lebesgue measure on $\mathbb{R}$ . . . 419
5.4.1 The Lebesgue outer measure and the Lebesgue measure on $\mathbb{R}$ . . . 419
5.4.2 Borel sets in $\mathbb{R}$ as examples of Lebesgue measurable sets . . . 423
5.4.3 Further properties of the Lebesgue measure on $\mathbb{R}$ . . . 428
5.4.4 Notes . . . 435

5.5 Lebesgue measure on $\mathbb{R}^n$ . . . 436
5.5.1 The Lebesgue outer measure and the Lebesgue measure on $\mathbb{R}^n$ . . . 436
5.5.2 Borel sets in $\mathbb{R}^n$ as examples of Lebesgue measurable sets . . . 438
5.5.3 Further properties of the Lebesgue measure on $\mathbb{R}^n$ . . . 442
5.5.4 Lebesgue measure on $\mathbb{R}^n$ as a product measure . . . 449
5.5.5 Coverings of subsets of $\mathbb{R}^n$ . . . 451
5.5.6 The Banach–Tarski Paradox . . . 455
5.5.7 Notes . . . 466
5.6 Measurable functions . . . 468
5.6.1 General measurable maps and functions . . . 468
5.6.2 Measurability and operations on functions . . . 474
5.6.3 Sequences of measurable functions . . . 479
5.6.4 $\mathbb{C}$- and vector-valued measurable functions . . . 486
5.6.5 Simple functions and approximations of measurable functions . . . 490
5.6.6 Topological characterisations of convergence for sequences of measurable functions . . . 493
5.7 Integration on measure spaces . . . 500
5.7.1 Definition of the integral . . . 500
5.7.2 The integral and operations on functions . . . 510
5.7.3 Limit theorems . . . 514
5.7.4 Integration with respect to probability measures . . . 520
5.7.5 Topological characterisations of limit theorems . . . 523
5.7.6 Image measure and integration by image measure . . . 525
5.7.7 The integral for $\mathbb{C}$- and vector-valued functions . . . 527
5.7.8 Integration with respect to signed, complex, and vector measures . . . 531
5.7.9 Notes . . . 542
5.8 Integration on products . . . 543
5.8.1 The product measure by integration . . . 543
5.8.2 The integral on product spaces . . . 546
5.8.3 Fubini's Theorem . . . 547
5.9 The single-variable Lebesgue integral . . . 556
5.9.1 Lebesgue measurable functions . . . 556
5.9.2 The (conditional) Lebesgue integral . . . 560
5.9.3 Properties of the Lebesgue integral . . . 563
5.9.4 Swapping operations with the Lebesgue integral . . . 569
5.9.5 Locally Lebesgue integrable functions . . . 572
5.9.6 Absolute continuity . . . 574
5.9.7 The Fundamental Theorem of Calculus for the Lebesgue integral . . . 581
5.9.8 Lebesgue points . . . 593
5.9.9 Maximal functions . . . 595
5.9.10 The change of variables formula . . . 595
5.9.11 Topological characterisations of the deficiencies of the Riemann integral 598
5.10 Notes . . . 601

Section 5.1

Some motivation for abstract measure theory and integration

In this section we illustrate the problems with the Riemann integral when it comes to dealing with spaces of integrable functions. We do this by first deriving a "measure theory," the Jordan measure, for Riemann integration, although this is not a theory of measure that satisfies the criterion we impose in our subsequent development of measure theory. What we shall see is that the difficulty arises from the fact that the Jordan measure only behaves well when one uses *finite* unions and intersections of sets. This leads to problems with sequential operations where there is an inherent need to be able to handle countable set theoretic unions and intersections. This is illustrated clearly in Example 5.1.10. We then illustrate why this phenomenon has repercussions for the Riemann integral. The problem, as we shall see, is that limits and Riemann integration do not commute; see Example 5.1.11.

Do I need to read this section? If you have already decided to read this chapter, and you do not already understand why it is necessary to move beyond the Riemann integral, then you should read this section. •

5.1.1 The Jordan measure and its limitations

We begin our discussion of the deficiencies of the Riemann integral by considering carefully the Jordan measure, which was touched lightly upon in Section **??**. Here we develop the Jordan measure in detail before finally tearing it down.

In Section **??** we introduced the idea of a Jordan measurable set as a set A whose characteristic function χ_A is Riemann integrable. In Theorem 5.1.5 we showed that a bounded set A is Jordan measurable if and only if $\mathrm{bd}(A)$ has zero volume if and only if $\mathrm{bd}(A)$ has zero measure. In this section we shall consider the Jordan measure in more detail and see that it has certain clear limitations.

First let us give a characterisation of Jordan measurable sets that will echo some of the constructions that will follow in our development of general measure theory. The basic building blocks for the Jordan measure are so-called elementary sets.

5.1.1 Definition (Elementary set) A subset $E \subseteq \mathbb{R}^n$ is ***elementary*** if $E = \cup_{j=1}^{k} C_j$ for bounded rectangles $C_1, \dots, C_k$. •

Note that, given a elementary set E, the expression of E as a union of bounded rectangles is not unique. Moreover, since there is no restriction that the rectangles do not overlap, the following result is of interest.

5.1.2 Proposition (Elementary sets are finite unions of disjoint rectangles) *If* E *is a elementary set then there exists disjoint rectangles* $\mathrm{C}_1, \dots, \mathrm{C}_\mathrm{k}$ *such that* $\mathrm{E} = \cup_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{C}_\mathrm{j}$.

Proof By definition we can write an elementary set as $E = \cup_{j=1}^{\tilde{k}} \tilde{C}_j$ for rectangles $\tilde{C}_1, \dots, \tilde{C}_{\tilde{k}}$. We shall prove the proposition by induction on $\tilde{k}$. The result is clearly true for $\tilde{k} = 1$. Suppose that the result is true for $\tilde{k} \in \{1, \dots, \tilde{m}\}$ and suppose that $E = \cup_{j=1}^{\tilde{m}+1} \tilde{C}_j$ and write

$$E = \left(\cup_{j=1}^{\tilde{m}} (\tilde{C}_j \cap \tilde{C}_{m+1})\right) \cup \left(\tilde{C}_{m+1} \setminus (\cup_{j=1}^{\tilde{m}} \tilde{C}_j))\right).$$

By the induction hypothesis there exists disjoint rectangles $C_1, \dots, C_l$ such that

$$\cup_{j=1}^{\tilde{m}} \tilde{C}_j = \cup_{j=1}^{l} C_j.$$

Thus

$$E = \left(\cup_{j=1}^{l} (C_j \cap \tilde{C}_{m+1})\right) \cup \left(\cup_{j=1}^{l} (\tilde{C}_{m+1} - C_j)\right).$$

Thus the result boils down to the following lemma.

1 Lemma *If* C *and* C′ *are bounded rectangles then* C∩C′ *is a bounded rectangle if it is nonempty and* C − C′ *is a finite union of disjoint bounded rectangles if it is nonempty.*

Proof Suppose that

$$C = I_1 \times \cdots \times I_n, \qquad C' = I'_1 \times \cdots \times I'_n$$

for bounded intervals $I_1, \dots, I_n$ and $I'_1, \dots, I'_n$. Note that $\boldsymbol{x} \in C \cap C'$ if and only if $x_j \in I_j \cap I'_j$, $j \in \{1, \dots, n\}$. That is,

$$C \cap C' = (I_1 \cap I'_1) \times \cdots \times (I_n \cap I'_n).$$

Since $(I_j \cap I'_j)$, $j \in \{1, \dots, n\}$, are bounded intervals if they are nonempty, it follows that $C \cap C'$ is a bounded rectangle if it is nonempty.

Note that $C - C' = C \setminus (C \cap C')$. We may as well suppose that each of the intersections $I_j \cap I'_j$, $j \in \{1, \dots, n\}$, is a nonempty bounded interval. Then write $I_j = J_j \cup (I_j \cap I'_j)$ where $J_j \cap (I_j \cap I'_j) = \emptyset$. This defines a partition of C where the interval I_j is partitioned as $(J_j, I_j \cap I'_j)$, $j \in \{1, \dots, n\}$. Thus this gives C as a finite disjoint union of rectangles, the subrectangles of the partition. Moreover, $C \cap C'$ corresponds exactly to the subrectangle

$$(I_1 \cap I'_1) \cap \cdots \cap (I_n \cap I'_n)$$

of this partition. By removing this subrectangle, we have $C - C'$ as a finite union of disjoint bounded rectangles, as desired. ▼

This completes the proof. ■

The previous result makes plausible the following definition.

5.1.3 Definition (Jordan measure of an elementary set) If $E \subseteq \mathbb{R}^n$ is an elementary set and if $E = \cup_{j=1}^{k} C_j$ for disjoint bounded rectangles $C_1, \dots, C_k$, then the ***Jordan measure*** of E is

$$\rho(E) = \sum_{j=1}^{k} \operatorname{vol}(C_j). \qquad \bullet$$

This definition has the possible ambiguity that it depends on writing E as a finite union of disjoint bounded rectangles, and such a union is not uniquely defined. However, one can refer to Proposition **??** to see that the definition is, in fact independent of how this union is made.

With the Jordan measure of elementary sets, we can introduce the following concepts which we shall see arise again when we are doing "serious" measure theory.

5.1.4 Definition (Inner and outer Jordan measure) If $A \subseteq \mathbb{R}^n$ is a bounded set then

(i) the ***Jordan outer measure*** of A is

$$\rho^*(A) = \inf\{\rho(E) \mid E \text{ an elementary set containing } A\}$$

and

(ii) the ***Jordan inner measure*** of A is

$$\rho_*(A) = \sup\{\rho(E) \mid E \text{ an elementary set contained in } A\} \qquad \bullet$$

Note that the Jordan outer and inner measures of a bounded set always exist, provided that, for the inner measure, we allow that the empty set be thought of as an elementary set, and that we adopt the (reasonable) convention that $\rho(\emptyset) = 0$.

The following result gives a characterisation of bounded Jordan measurable sets, including some of the characterisations we have already proved in Section **??**.

5.1.5 Theorem (Characterisations of bounded Jordan measurable sets) *For a bounded subset* $A \subseteq \mathbb{R}^n$ *the following statements are equivalent:*

(i) A *is Jordan measurable;*

(ii) $\mathrm{vol}(\mathrm{bd}(A)) = 0$*;*

(iii) χ_A *is Riemann integrable;*

(iv) $\rho^*(A) = \rho_*(A)$.

Proof The equivalent of the first three statements is the content of Theorems 5.1.5 and **??**. Thus we only prove the equivalence of the last statement with the other three.

Let C be a fat compact rectangle containing A.

First suppose that A is Jordan measurable and let $\epsilon \in \mathbb{R}_{>0}$. Since χ_A is Riemann integrable there exists a partition $\boldsymbol{P}$ of C such that

$$A_+(\chi_A, \boldsymbol{P}) - A_-(\chi_A, \boldsymbol{P}) < \epsilon.$$

Let the subrectangles of $\boldsymbol{P}$ be divided into three sorts: (1) the first sort are those subrectangles that lie within A; (2) the second sort are those that intersect A; (3) the third sort are rectangles that do not intersect A. From the definition of χ_A, $A_+(\chi_A, \boldsymbol{P})$ is the total volume of the rectangles of the third sort and $A_-(\chi_A, \boldsymbol{P})$ is the total volume of the rectangles of the first sort. Moreover, by the definitions of these rectangles,

$$\rho^*(A) \le A_+(\chi_A, \boldsymbol{P}), \quad \rho_*(A) \ge A_-(\chi_A, \boldsymbol{P}).$$

Thus $\rho^*(A) - \rho_*(A) < \epsilon$, giving $\rho^*(A) = \rho_*(A)$ since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary.

Now suppose that $\rho^*(A) = \rho_*(A)$, let $\epsilon \in \mathbb{R}_{>0}$, and let $\overline{E}_\epsilon$ and $\underline{E}_\epsilon$ be elementary subsets of $\mathbb{R}^n$ such that $\rho(\overline{E}_\epsilon) - \rho(\underline{E}_\epsilon) < \epsilon$. Since $\overline{E}_\epsilon$ is a disjoint union of finitely many bounded rectangles there exists a partition $\overline{P}_\epsilon$ of C such that $\overline{E}_\epsilon$ is a union of subrectangles from $\overline{\boldsymbol{P}}_\epsilon$. Similarly, there exists a partition $\underline{\boldsymbol{P}}_\epsilon$ such that $\underline{E}_\epsilon$ is a union of subrectangles of $\underline{\boldsymbol{P}}_\epsilon$. Now let $\boldsymbol{P}_\epsilon$ be a partition that refines both $\overline{\boldsymbol{P}}_\epsilon$ and $\underline{\boldsymbol{P}}_\epsilon$. Then we have

$$A_+(\chi_A, \boldsymbol{P}_\epsilon) \le \rho^*(\overline{E}_\epsilon), \quad A_-(\chi_A, \boldsymbol{P}_\epsilon \ge \rho_*(\underline{E}_\epsilon),$$

which gives

$$A_+(\chi_A, \boldsymbol{P}_\epsilon) - A_-(\chi_A, \boldsymbol{P}_\epsilon) < \epsilon,$$

as desired. ■

Note that it is only the basic definition of a Jordan measurable set, i.e., that its boundary have measure zero, that is applicable to unbounded sets. However, we can still use the characterisation of bounded Jordan measurable sets to give the measure of possibly unbounded sets. For the following definition we denote by

$$C_R = [-R, R] \times \cdots \times [-R, R]$$

the rectangle centred at $\mathbf{0}$ whose sides have length $2R$ for $R \in \mathbb{R}_{>0}$.

5.1.6 Definition (Jordan measure[1]) Let $\mathscr{J}(\mathbb{R}^n)$ denote the collection of Jordan measurable sets of $\mathbb{R}^n$ and define $\rho\colon \mathscr{J}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\rho(A) = \lim_{R\to\infty} \rho^*(A \cap C_R),$$

noting that $A \cap C_R$ is a bounded Jordan measurable set for each $R \in \mathbb{R}_{>0}$. For $A \in \mathscr{J}(\mathbb{R}^n)$, $\rho(A)$ is the ***Jordan measure*** of A. •

Of course, by Theorem 5.1.5 we could as well have defined

$$\rho(A) = \lim_{R\to\infty} \rho_*(A \cap C_R).$$

Let us look at some examples that flesh out the definition.

5.1.7 Examples (Jordan measurable sets)

1. $\mathbb{R}^n$ is itself Jordan measurable and $\rho(\mathbb{R}^n) = \infty$.
2. Let us consider the set

   $$A = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_2| \leq \mathrm{e}^{-|x_1|}\}$$

   An application of Fubini's Theorem gives

   $$\int_A \mathrm{d}\boldsymbol{x} = \int_{-\infty}^{\infty} \Big(\int_{-\mathrm{e}^{-|x_1|}}^{\mathrm{e}^{-|x_1|}} \mathrm{d}x_2 \Big) \mathrm{d}x_1 = 4.$$

   By the definition of the Riemann integral for unbounded domains (see Definition **??**) this means that $\rho(A) = 4$. Thus unbounded domains can have finite Jordan measure. •

The following property of Jordan measures—or more precisely the fact that *only* the following result applies—is crucial to why they are actually not so useful.

[1] The Jordan measure is *not* a measure as we shall define the notion in Section 5.3. However, it is convenient to write as if it is to get prepared for the more general and abstract development to follow.

5.1.8 Proposition (Jordan measurable sets are closed under finite intersections and unions) *If* $\mathrm{A}_1, \dots, \mathrm{A}_\mathrm{k} \in \mathscr{J}(\mathbb{R}^\mathrm{n})$ *are Jordan measurable then* $\cap_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{A}_\mathrm{j}, \cup_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{A}_\mathrm{j} \in \mathscr{J}(\mathbb{R}^\mathrm{n})$.

Proof This is straightforward and we leave the details to the reader as Exercise 5.1.1. ■

Having now built up the Jordan measure and given some of its useful properties, let us now proceed to show that it has some very undesirable properties. This destruction of the Jordan measure is tightly connected with our bringing down of the Riemann integral in the next section. Sometimes, in order to understand why something is useful (in this case, the Lebesgue measure), it helps to first understand why the alternatives are *not* useful. It is with this in mind that the reader should undertake to read the remainder of this section.

The most salient question about the Jordan measure is, "What are the Jordan measurable sets?" The first thing we shall note is that there are "nice" open sets that are not Jordan measurable. This is not good, since open sets form the building blocks of the topology of $\mathbb{R}^n$.

5.1.9 Example (A regularly open non-Jordan measurable set) We shall construct a subset $A \subseteq [0, 1]$ with the following properties:

1. A is open;
2. $A = \operatorname{int}(\operatorname{cl}(A))$ (an open set with this property is called ***regularly open***);
3. A is not Jordan measurable.

The construction is involved, and will be presented with the aid of a series of lemmata. If you are prepared to take the existence of a set A as stated on faith, you can skip the details. Let us denote $I = [0, 1]$.

Any $x \in I$ can be written in the form

$$\sum_{j=1}^{\infty} \frac{a_j}{3^j}$$

for $a_j \in \{0, 1, 2\}$. This is called a ***ternary decimal expansion*** of x, and we refer the reader to Exercise 2.4.8 for details of this construction in base 10. There is a possible nonuniqueness in such decimal expansions that arises in the following manner. If $a_1, \dots, a_k \in \{0, 1, 2\}$ then the numbers

$$\sum_{j=1}^{k} \frac{a_j}{3^j} + \sum_{j=k+1}^{\infty} \frac{2}{3^j} \qquad \text{and} \qquad \sum_{j=1}^{k-1} \frac{a_j}{3^j} + \frac{(a_k + 1) \mod 3}{3^k} + \sum_{j=k+1}^{\infty} \frac{0}{3^k}$$

are the same, where

$$(a_k + 1) \mod 3 = \begin{cases} a_k + 1, & a_k \in \{0, 1\}, \\ 0, & a_k = 2. \end{cases}$$

Now, for $k \in \mathbb{Z}_{>0}$, define B_k to be the subset of I for which, if $x \in B_k$ is written as

$$x = \sum_{j=1}^{\infty} \frac{a_j}{3^j},$$

then $a_j = 1$ for $j \in \{2^{k-1}+1, 2^{k-1}+2, \dots, 2^k\}$. For numbers with nonunique ternary decimal expansions, we ask that *both* representations satisfy the condition.

1 **Lemma** *For* $\mathrm{k} \in \mathbb{Z}_{>0}$*,* $\mathrm{B_k}$ *is a disjoint union of* $3^{2^{\mathrm{k}-1}}$ *open intervals each of length* $\frac{1}{3^{2^{\mathrm{k}}}}$.

Proof For $\boldsymbol{a} = (a_1, \dots, a_{2^{k-1}}) \in \{0,1,2\}$ define $I_{\boldsymbol{a}}$ to be the open interval whose left endpoint is

$$\sum_{j=1}^{2^{k-1}} \frac{a_j}{3^j} + \sum_{j=2^{k-1}+1}^{2^k} \frac{1}{3^j}$$

and whose right endpoint is

$$\sum_{j=1}^{2^{k-1}} \frac{a_j}{3^j} + \sum_{j=2^{k-1}+1}^{2^k} \frac{1}{3^j} + \sum_{j=2^k+1}^{\infty} \frac{2}{3^j}.$$

There are obviously $3^{2^{k-1}}$ such intervals and each such interval has length 3^{2^k}. One can directly verify that B_k is the union of all of these intervals. ▼

Now define $B = \cup_{k=1}^{\infty} B_k$ which is, therefore, open. The sets B_k, $k \in \mathbb{Z}_{>0}$, satisfy the following.

2 **Lemma** *If* $\mathrm{l}, \mathrm{k} \in \mathbb{Z}_{>0}$ *satisfy* $\mathrm{l} < \mathrm{k}$ *then* $\mathrm{bd}(\mathrm{B_l}) \cap \mathrm{B_k} = \emptyset$.

Proof Let $x \in \mathrm{bd}(B_l)$. Then

$$x = \sum_{j=1}^{\infty} \frac{a_j}{3^j}$$

where either $a_j = 0$ for all $j \geq 2^l$ or $a_j = 2$ for all $j \geq 2^l$. Thus $a_{2^k} \neq 1$ and so $x \notin B_k$. ▼

Now, for $k \in \mathbb{Z}_{>0}$, we define

$$A_k = B_k - \big(\mathrm{cl}(B_{k+1}) \cup \big(\cup_{j=1}^{k-1} B_j\big)\big).$$

These sets have the following property.

3 **Lemma** $\mathrm{A_k} = \mathrm{B_k} \cap (\mathrm{I} \setminus \mathrm{cl}(\mathrm{B_{k+1}})) \cap_{\mathrm{j}=1}^{\mathrm{k}-1} (\mathrm{I} \setminus \mathrm{cl}(\mathrm{B_j}))$*. In particular,* $\mathrm{A_k}$ *is open for each* $\mathrm{k} \in \mathbb{Z}_{>0}$.

Proof By DeMorgan's Laws we have

$$A_k = B_k \cap (I \setminus \mathrm{cl}(B_{k+1})) \cap_{j=1}^{k-1} (I \setminus B_j).$$

By Lemma 2 we have

$$B_k \cap (I \setminus B_j) = B_k \cap (I \setminus \mathrm{cl}(B_j))$$

for each $j \in \{1, \dots, k-1\}$, and the stated formula for A_k follows from this. That A_k is open follows since finite intersections of open sets are open. ▼

Thus the set $A = \cup_{k=1}^{\infty} A_k$ is open, being a union of open sets, and is contained in B since $A_k \subseteq B_k$ for each $k \in \mathbb{Z}_{>0}$.

Now, for $k \in \mathbb{Z}_{>0}$, define

$$C_k = (B_k \cap B_{k+1}) \setminus \left(\cup_{j=1}^{k-1} B_j\right).$$

By the same argument as employed in the proof of Lemma 3, Lemma 2 implies that

$$C_k = B_k \cap B_{k+1} \cap_{j=1}^{k-1} (I \setminus \mathrm{cl}(B_j))$$

and so C_k, $k \in \mathbb{Z}_{>0}$, is open, being a finite intersection of open sets. Then define the open set $C = \cup_{k=1}^{\infty} C_k$. The relationship between the sets A_l, $l \in \mathbb{Z}_{>0}$, and C_k, $k \in \mathbb{Z}_{>0}$.

4 **Lemma** *For each* $l, k \in \mathbb{Z}_{>0}$, $A_l \cap C_k = \emptyset$.

Proof First suppose that $l = k$. By definition we have

$$A_k \subseteq I \cap \mathrm{cl}(B_{k+1}), \quad C_k \subseteq B_{k+1}$$

which immediately gives $A_k \cap C_k = \emptyset$. Now suppose that $l < k$. Again by definition we have

$$A_l \subseteq B_l, \quad C_k \subseteq I \setminus \mathrm{cl}(B_l),$$

giving $A_l \cap C_k = \emptyset$. Finally, for $l > k$ we have

$$A_l \subseteq I \setminus \mathrm{cl}(B_k), \quad C_k \subseteq B_k,$$

giving $A_l \cap C_k = \emptyset$. ▼

The following lemma then gives a relationship between A and C.

5 **Lemma** $\mathrm{cl}(A) = I \setminus C$.

Proof By Lemma 4 we have $A \cap C = \emptyset$. That is, $A \subseteq I \setminus C$. Since $I \setminus C$ is closed it follows that $\mathrm{cl}(A) \subseteq I \setminus C$. The difficult bit if the converse inclusion. Let $x \in I \setminus C$. We consider three cases.

1. $x \in \cup_{k=1}^{\infty} \mathrm{bd}(A_k)$: In this case, since $\mathrm{bd}(A_k) \subseteq \mathrm{cl}(A_k) \subseteq \mathrm{cl}(A)$ for each $k \in \mathbb{Z}_{>0}$ it immediately follows that $x \in \mathrm{cl}(A)$.
2. $x \notin B$: In this case we can write

$$x = \sum_{j=1}^{\infty} \frac{a_j}{3^j}.$$

Since $x \notin B$, for every $k \in \mathbb{Z}_{>0}$ there exists $j \in \{2^{k-1}+1, \dots, 2^k\}$ such that $a_j \neq 1$. Now define a sequence $(y_k)_{k \in \mathbb{Z}_{>0}}$ by asking that $y_k = \sum_{j=1}^{\infty} \frac{b_j}{3^j}$ with

$$b_j = \begin{cases} 1, & j \in \{2^{k-1}+1, \dots, 2^k\}, \\ a_j, & \text{otherwise}. \end{cases}$$

We then have $|x - y_k| \leq \frac{1}{3^{2^k}}$ (cf. the proof of Lemma 1) and so the sequence $(y_k)_{k \in \mathbb{Z}_{>0}}$ converges to x. Moreover, by construction,

$$y_k \in B_k,\ y_k \notin B_1 \cup \cdots \cup B_{k-1},\ y_k \notin \mathrm{cl}(B_{k+1}).$$

(Only the last of these statements is potentially not obvious. It, however, follows from the characterisation of B_{k+1}, and by implication the characterisation of $\mathrm{cl}(B_{k+1})$, obtained in Lemma 1.) That is, by definition of A_k, $y_k \in A_k \subseteq A$. Thus $x \in \mathrm{cl}(A)$ by Proposition 2.5.18.

3. $x \notin \cup_{k=1}^{\infty} \mathrm{bd}(A_k)$ and $x \in B$: Let $k \in \mathbb{Z}_{>0}$ be the least index for which $x \in B_k$. Since $x \notin C$ it follows that $x \notin C_k$ and so $x \notin B_{k+1}$ and $x \notin B_j$ for $j \in \{1, \dots, k-1\}$. We also have $x \notin \mathrm{bd}(A_{k+1})$. We claim that $\mathrm{bd}(B_{k+1}) \subseteq \mathrm{bd}(A_{k+1})$. Indeed, for each $m \in \mathbb{Z}_{>0}$, by construction of the set A_m, $\mathrm{bd}(A_m)$ consists of those ternary decimal expansions $\sum_{j=1}^{\infty} \frac{a_j}{3^j}$ having the following three properties:
   (a) for $l < m$ there exists $j \in \{2^{l-1}+1, \dots, 2^l\}$ such that $a_j \neq 1$;
   (b) $a_j = 1$ for each $j \in \{2^{m-1}+1, \dots, 2^m\}$;
   (c) there exists $j \in \{2^m+1, \dots, 2^{m+1}\}$ such that $a_j \neq 1$.

   Using this characterisation, and by referring to the description of B_m in Lemma 1, we then see that, indeed, $\mathrm{bd}(B_{k+1}) \subseteq \mathrm{bd}(A_{k+1})$. Thus we conclude that $x \notin \mathrm{bd}(B_{k+1})$. Then, by definition of A_k, $x \in A_k \subseteq A \subseteq \mathrm{cl}(A)$, as desired. ▼

We also then have

$$\mathrm{int}(\mathrm{cl}(A)) \subseteq \mathrm{int}(I \setminus C) = I \setminus \mathrm{cl}(C).$$

That is, $\mathrm{int}(\mathrm{cl}(A)) \cap \mathrm{cl}(C) = \emptyset$.

Now we can prove that A has one of the properties we set out for it to have.

6 **Lemma** A = int(cl(A)).

Proof Since $A \subseteq \mathrm{cl}(A)$ we have $A = \mathrm{int}(A) \subseteq \mathrm{int}(\mathrm{cl}(A))$. It is thus the converse inclusion we must prove.

We first claim that $\mathrm{int}(\mathrm{cl}(A)) \subseteq B$. Suppose that $x \notin B$. Let us write

$$x = \sum_{j=1}^{\infty} \frac{a_j}{3^j}.$$

Since $x \notin B$, for every $k \in \mathbb{Z}_{>0}$ there exists $j \in \{2^{k-1}+1, \dots, 2^k\}$ such that $a_j \neq 1$. Now, for $k \in \mathbb{Z}_{>0}$, define

$$y_k = \sum_{j=1}^{\infty} \frac{c_j}{3^j}$$

where

$$c_j = \begin{cases} a_j, & j \leq 2^{k-1}, \\ 1, & j > 2^{k-1}. \end{cases}$$

Then one can directly verify that

$$y_k \in B_k,\ y_k \in B_{k+1},\ y_k \notin B_1 \cup \cdots \cup B_{k-1}.$$

Thus, by definition of C_k, $y \in C_k$. Moreover, $|x - y_k| \leq \frac{1}{3^{2^{k-1}}}$ and so the sequence $(y_k)_{k\in\mathbb{Z}_{>0}}$ converges to x. Therefore, since $y_k \in C$ for each $k \in \mathbb{Z}_{>0}$, $x \in \mathrm{cl}(C)$ by Proposition 2.5.18. Thus $x \notin \mathrm{int}(\mathrm{cl}(A))$ by our computation just preceding the statement of the lemma.

Now, if $x \in \mathrm{int}(\mathrm{cl}(A))$ then $x \in B$ and we let $k \in \mathbb{Z}_{>0}$ be the least integer for which $x \in B_k$. We claim that $x \notin \mathrm{cl}(B_{k+1})$. We suppose that $x \in \mathrm{cl}(B_{k+1})$ and arrive at a contradiction. There are two possibilities.

1. $x \in \mathrm{bd}(B_{k+1})$: First of all, using the characterisation of the sets B_l, $l \in \mathbb{Z}_{>0}$, from Lemma 1 and using the definition of the sets C_l, $l \in \mathbb{Z}_{>0}$, we deduce that $\mathrm{bd}(B_l) \subseteq \mathrm{bd}(C_l)$ for each $l \in \mathbb{Z}_{>0}$. Therefore, if $x \in \mathrm{bd}(B_{k+1})$ then $x \in \mathrm{bd}(C_{k+1}) \subseteq \mathrm{cl}(C_{k+1}) \subseteq \mathrm{cl}(C)$. This contradicts the fact that $x \in \mathrm{int}(\mathrm{cl}(A))$ and that $\mathrm{int}(\mathrm{cl}(A)) \cap \mathrm{cl}(C) = \emptyset$.
2. $x \in B_{k+1}$: In this case $x \in B_k \cap B_{k+1} \subseteq C_k \subseteq \mathrm{cl}(C)$, and we arrive at a contradiction, just as in the previous case.

Thus we have shown that $x \notin \mathrm{cl}(B_{k+1})$. But, by definition, this implies that $x \in A_k \subseteq A$, since $x \notin \cup_{j=1}^{k-1} B_j$ by definition of k. ▼

Finally, to complete the example, we need only show that A is not Jordan measurable. To do this, we shall show that $\mathrm{bd}(A)$ does not have measure zero. In fact, we shall show that $\mathrm{bd}(A)$ has positive measure, but this relies on actually knowing what "measure" means; it means Lebesgue measure. We shall subsequently carefully define Lebesgue measure, but all we need to know here is that (1) the Lebesgue measure of a countable collection of intervals is less than or equal to the sum of the lengths of the intervals and (2) the Lebesgue measure of two disjoint sets is the sum of their measures. Let us denote by $\lambda(S)$ the Lebesgue measure of a set S. We note that, by Lemma 1,

$$\lambda(B_k) = 3^{2^{k-1}} \frac{1}{3^{2^k}} = \frac{1}{3^{2^{k-1}}}.$$

Thus

$$\lambda(B) \leq \sum_{k=1}^{\infty} \frac{1}{3^{2^{k-1}}} < \sum_{j=1}^{\infty} \frac{1}{3^j} = \frac{1}{2}$$

(how would you compute this sum?). Since $A \subseteq B$ we also have $\lambda(A) < \frac{1}{2}$. Therefore, since $\mathrm{cl}(A) = I \setminus C$ and since $C \subseteq B$,

$$\lambda(A) + \lambda(\mathrm{bd}(A)) = \lambda(\mathrm{cl}(A)) \geq \lambda(I \setminus B) = 1 - \lambda(B) > \tfrac{1}{2} > \lambda(A),$$

which gives $\lambda(\mathrm{bd}(A)) \in \mathbb{R}_{>0}$, so A is not Jordan measurable. •

This is a rather complicated example. However, it says something important. It says that not all open sets are Jordan measurable, not even "nice" open sets (and regularly open sets are thought of as being pretty darn nice). Open subsets of $\mathbb{R}$ are pretty easy to describe. Indeed, by Proposition 2.5.6 such sets are countable unions of open intervals. If one has an open subset of $[0, 1]$, such as the one just constructed, this means that the total lengths of these intervals should sum to a finite number of value at most one. This should, if the world is right, be the "measure" of this open set. However, the example indicates that this is just not so if "measure" means "Jordan measure." We shall see that it *is* so for the Lebesgue measure.

In Proposition 5.1.8 we stated that finite unions and intersections of Jordan measurable sets are Jordan measurable. This no longer holds if one replaces "finite" with "countable."

5.1.10 Examples (Jordan measurable sets are not closed under countable intersections and unions)

1. Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be an enumeration of the rational numbers in the interval $[0, 1]$. For each $j \in \mathbb{Z}_{>0}$ the set $\{q_j\}$ is Jordan measurable with Jordan measure 0. Thus, by Proposition 5.1.8 any finite union of these sets is also Jordan measurable with Jordan measure 0. However, the set $\cup_{j=1}^{\infty}\{q_j\}$ is not Jordan measurable by Example 3.4.10.
2. Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be as above and define $A_j = [0, 1] \setminus \{q_j\}$. Then A_j is Jordan measurable and has Jordan measure 1. Moreover, any finite intersection of these sets is Jordan measurable with Jordan measure 1. However, $\cap_{j=1}^{\infty} A_j$ is equal to the set of irrational numbers in the interval $[0, 1]$ and is not Jordan measurable in exactly the same manner as the set $\cup_{j=1}^{\infty}\{q_j\}$ is not Jordan measurable, cf. Example 3.4.10. •

A good question is, "Who cares if the Jordan measure is not closed under countable intersections and unions?" This is not obvious, but it certainly underlies, for example, the failure of the set in Example 5.1.9 to be Jordan measurable. Somewhat more precisely, this failure of the Jordan measure to not be closed under countable set theoretic operations is the reason why the Riemann integral does not have nice properties with respect to sequences, as we now explain explicitly.

5.1.2 Some limitations of the Riemann integral

In this section we simply give an example that illustrates a fundamental defect with the theory of Riemann integration. The problem we illustrate is the lack of commutativity of limits and Riemann integration. The reader may wish to refer to the discussion in Section **??** concerning the Monotone and Dominated Convergence Theorems for the Riemann integral to get more insight into this.

5.1.11 Example (Limits do not commute with Riemann integration) First recall from Example 3.4.10 that the function $f: [0, 1] \to \mathbb{R}$ defined as taking value 1 on rational numbers, and value 0 on irrational numbers is not Riemann integrable. It is legitimate to inquire why one should care if such a degenerate function should be integrable. The reason is that the function f arises as the limit of a sequence of integrable functions. We explain this in the following example.

By Exercise 2.1.3, the set of rational numbers in $[0,1]$ is countable. Thus it is possible to write the set of rational numbers as $(q_j)_{j\in\mathbb{Z}_{>0}}$. For each $j\in\mathbb{Z}_{>0}$ define $f_j\colon[0,1]\to\mathbb{R}$ by

$$f_j(x)=\begin{cases}1, & x=q_j,\\ 0, & \text{otherwise.}\end{cases}$$

One may readily verify that f_j is Riemann integrable for each $j\in\mathbb{Z}_{>0}$, and that the value of the Riemann integral is zero. By Proposition 3.4.22 it follows that for $k\in\mathbb{Z}_{>0}$, the function

$$g_k=\sum_{j=1}^{k}f_j$$

is Riemann integrable, and that the value of the Riemann integral is zero. Thus we have

$$\lim_{k\to\infty}g_k(x)=f(x),\quad \lim_{k\to\infty}\int_a^b g_k(x)\,\mathrm{d}x=0,$$

the left limit holding for each $x\in[0,1]$ (i.e., the sequence $(g_k)_{k\in\mathbb{Z}_{>0}}$ converges pointwise to f). It now follows that

$$\lim_{k\to\infty}\int_a^b g_k(x)\,\mathrm{d}x\neq\int_a^b\lim_{k\to\infty}g_n(x)\,\mathrm{d}x.$$

Indeed, the expression on the right hand side is not even defined! •

It is perhaps not evident immediately why this lack of commutativity of limits and integrals is in any way debilitating, particularly given the inherent silliness of the functions in the preceding example. We shall not really understand the reasons for this in any depth until we consider in detail convergence theorems in Section 5.7.3.

Let us illustrate some additional "features" of the Riemann integral, the exact context for which we will only consider in detail in Chapter 6 (see, in particular, Sections 6.7.7 and 6.7.8). We shall freely use the language and notation from that chapter. Let us define

$$\mathsf{R}^{(1)}([0,1];\mathbb{R})=\{f\colon[0,1]\to\mathbb{R}\mid f\text{ is Riemann integrable}\},$$

and recall from Propositions 3.4.22 and 3.4.25 that $\mathsf{R}^{(1)}([0,1];\mathbb{R})$ is a $\mathbb{R}$-vector space. Now let us define a seminorm $\|\cdot\|_1$ on $\mathsf{R}^{(1)}([0,1];\mathbb{R})$ by

$$\|f\|_1=\int_0^1|f(x)|\,\mathrm{d}x.$$

This fails to be a norm because there exist nonzero Riemann integrable functions f on $[0,1]$ for which $\|f\|_1=0$ (for example, take f to be a function that has a nonzero value at a single point in $[0,1]$). To produce a normed vector space we denote

$$Z([0,1];\mathbb{R})=\{f\in\mathsf{R}^{(1)}([0,1];\mathbb{R})\mid \|f\|_1=0\},$$

and by Theorem 6.1.8 note that

$$\mathsf{R}^1([0,1];\mathbb{R}) \triangleq \mathsf{R}^{(1)}([0,1];\mathbb{R})/Z([0,1];\mathbb{R})$$

is a normed vector space when equipped with the norm

$$\|f + Z([0,1];\mathbb{R})\|_1 \triangleq \|f\|_1,$$

where we use the abuse of notation of using the same symbol $\|\cdot\|_1$ for the norm. Note that $\mathsf{R}^1([0,1];\mathbb{R})$ is a vector space, not of functions, but of equivalence classes of functions under the equivalence relation that two Riemann integrable functions are equivalent when the absolute value of their difference has zero integral.

The crux of the matter is now the following result, the proof of which makes free use of concepts in this chapter that we have not yet introduced.

5.1.12 Proposition (The normed vector space of Riemann integrable functions is not complete) *The $\mathbb{R}$-normed vector space $(\mathsf{R}^1([0,1];\mathbb{R}), \|\cdot\|_1)$ is not complete.*

Proof Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be an enumeration of the rational numbers in $[0,1]$. Let $\ell \in (0,1)$ and for $j \in \mathbb{Z}_{>0}$ define

$$I_j = [0,1] \cap (q_j - \tfrac{\ell}{2^{j+1}}, q_j + \tfrac{\ell}{2^{j+1}})$$

to be the interval of length $\frac{\ell}{2^j}$ centred at q_j. Then define $A_k = \cup_{j=1}^k I_j$, $k \in \mathbb{Z}_{>0}$, and $A = \cup_{j\in\mathbb{Z}_{>0}} A_j$. Also define $f_k = \chi_{A_k}$, $k \in \mathbb{Z}_{>0}$, and $f = \chi_A$ be the characteristic functions of A_k and A, respectively. Note that A_k is a union of a finite number of intervals and so f_k is Riemann integrable for each $k \in \mathbb{Z}_{>0}$. However, we claim that f is not Riemann integrable. Indeed, the characteristic function of a set is Riemann integrable if and only the boundary of the set has measure zero; this is a direct consequence of Lebesgue's theorem stating that a function is Riemann integrable if and only if its set of discontinuities has measure zero (Theorem 3.4.11). Note that since $\mathrm{cl}(\mathbb{Q}\cap[0,1]) = [0,1]$ we have

$$[0,1] = \mathrm{cl}(A) = A \cup \mathrm{bd}(A).$$

Thus

$$\lambda([0,1]) \leq \lambda(A) + \lambda(\mathrm{bd}(A)).$$

Since

$$\lambda(A) \leq \sum_{j=1}^{\infty} \lambda(I_j) \leq \ell,$$

it follows that $\lambda(\mathrm{bd}(A)) \geq 1 - \ell \in \mathbb{R}_{>0}$. Thus f is not Riemann integrable, as claimed.

Next we show that if $g\colon [0,1] \to \mathbb{R}$ satisfies $[g] = [f]$, then g is not Riemann integrable. To show this, it suffices to show that g is discontinuous on a set of positive measure. We shall show that g is discontinuous on the set $g^{-1}(0) \cap \mathrm{bd}(A)$. Indeed, let $x \in g^{-1}(0) \cap \mathrm{bd}(A)$. Then, for any $\epsilon \in \mathbb{R}_{>0}$ we have $(x-\epsilon, x+\epsilon) \cap A \neq \emptyset$ since $x \in \mathrm{bd}(A)$. Since $(x-\epsilon, x+\epsilon) \cap A$ is a nonempty open set, it has positive measure. Therefore, since f and g agree almost everywhere, there exists $y \in (x-\epsilon, x+\epsilon) \cap A$ such that $g(y) = 1$. Since this holds for every $\epsilon \in \mathbb{R}_{>0}$ and since $g(x) = 0$, it follows that g is discontinuous at x. Finally, it suffices to show that $g^{-1}(0) \cap \mathrm{bd}(A)$ has positive measure. But this follows since $\mathrm{bd}(A) = f^{-1}(0)$ has positive measure and since f and g agree almost everywhere.

We claim that the sequence $([f_k])_{k\in\mathbb{Z}_{>0}}$ is Cauchy in $\mathsf{R}^1([0,1];\mathbb{R})$. Let $\epsilon\in\mathbb{R}_{>0}$. Note that $\sum_{j=1}^{\infty}\lambda(I_j)\leq\ell$. This implies that there exists $N\in\mathbb{Z}_{>0}$ such that $\sum_{j=k+1}^{m}\lambda(I_j)<\epsilon$ for all $k,m\geq N$. Now note that for $k,m\in\mathbb{Z}_{>0}$ with $m>k$, the functions f_k and f_m agree except on a subset of $I_{k+1}\cup\cdots\cup I_m$. On this subset, f_m has value 1 and f_k has value 0. Thus

$$\int_0^1|f_m(x)-f_k(x)|\,\mathrm{d}x\leq\lambda(I_{k+1}\cup\cdots\cup I_m)\leq\sum_{j=k+1}^{m}\lambda(I_j).$$

Thus we can choose $N\in\mathbb{Z}_{>0}$ sufficiently large that $\|f_m-f_k\|_1<\epsilon$ for $k,m\geq N$. Thus the sequence $([f_k])_{k\in\mathbb{Z}_{>0}}$ is Cauchy, as claimed.

We next show that the sequence $([f_k])_{k\in\mathbb{Z}_{>0}}$ converges to $[f]$ in $\mathsf{L}^1([0,1];\mathbb{R})$ (see Section 6.7.7). Since the sequence $([f-f_k])_{k\in\mathbb{Z}_{>0}}$ is in the subset

$$\{[f]\in\mathsf{L}^1([0,1];\mathbb{R})\mid\ |f(x)|\leq 1\text{ for almost every }x\in[0,1]\},$$

by the Dominated Convergence Theorem, Theorem 5.7.28, it follows that

$$\lim_{k\to\infty}\|f-f_k\|_1=\int_I\lim_{k\to\infty}|f-f_k|\,\mathrm{d}\lambda=0.$$

This gives us the desired convergence of $([f_k])_{k\in\mathbb{Z}_{>0}}$ to $[f]$ in $\mathsf{L}^1([0,1];\mathbb{R})$. However, above we showed that $[f]\notin\mathsf{R}^1([0,1];\mathbb{R})$. Thus the Cauchy sequence $([f_k])_{k\in\mathbb{Z}_{>0}}$ in $\mathsf{R}^1([0,1];\mathbb{R})$ is not convergent in $\mathsf{R}^1([0,1];\mathbb{R})$, giving the desired incompleteness of $(\mathsf{R}^1([0,1];\mathbb{R}),\|\cdot\|_1)$. ■

It should be emphasised that all of the above "problems" are not so much one with using the Riemann integral to compute the integral of a given function, as to use the notion of a Riemann integrable function in stating theorems, particularly those where limits are involved. This problem is taken care of by the Lebesgue integral, to which we turn our attention in Section 5.7.1 in a general setting for integration.

5.1.3 An heuristic introduction to the Lebesgue integral

Before we get to the powerful general theory, we provide in this section an alternate way of thinking about the integral of a function defined on a compact interval. The idea is an essentially simple one. One defines the Riemann integral by taking increasingly finer partitions of the independent variable axis, where on each subinterval of the partition the approximation is constant. For the Lebesgue integral, it turns out that what one should do instead is partition the *dependent* variable axis.

The reader should not treat the following discussion as the definition of the Lebesgue integral. This definition will be provided precisely in the general framework of Section 5.7.1. But let us be a little precise about the idea. We let $I=[a,b]$ and let $f\colon I\to\mathbb{R}$ be a positive bounded function. This means that $f(I)\subset[0,M]$ for some $M\in\mathbb{R}_{>0}$. We then let P be a partition of $[0,M]$ with endpoints $(y_0=0,y_1,\dots,y_{n-1},y_n=M)$. Corresponding to this partition let us define sets

$$A_j=\{x\in I\mid\ f(x)\in[y_{j-1},y_j)\},$$

and then define

$$f_P = \sum_{j=1}^{n} y_j \chi_{A_j}.$$

The function f_P is called a ***simple function***, as we shall see in Section 5.7, and approximates f from below as depicted in Figure 5.1. The integral of one of these approximations is then

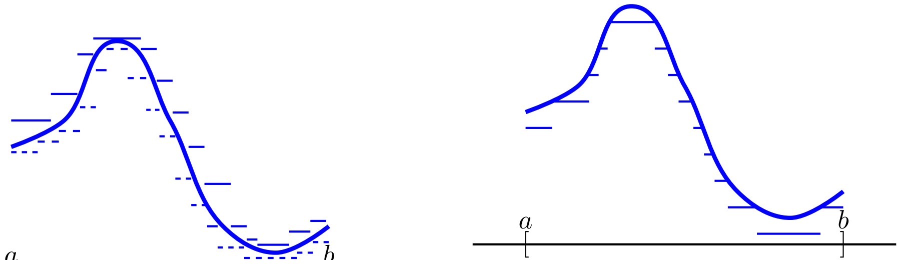


Figure 5.1 The idea behind the Riemann integral (left) and the Lebesgue integral (right)

$$\int_a^b f_P(x)\,\mathrm{d}x = \sum_{j=1}^{n} y_j \lambda(A_j),$$

where $\mu(A_j)$ is the "size" of the set A_j. If A_j is a union on intervals, then $\mu(A_j)$ is the sum of the lengths of these intervals. More generally, we shall define

$$\lambda(A) = \inf\Big\{ \sum_{j=1}^{\infty} |b_j - a_j| \;\Big|\; A \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j) \Big\}$$

for a very general class of subsets of $\mathbb{R}$. To define the integral of f we take

$$\text{“}\int_a^b f(x)\,\mathrm{d}x\text{”} = \sup\Big\{ \int_a^b f_P(x)\,\mathrm{d}x \;\Big|\; P \text{ a partition of } [0, M] \Big\}.$$

The idea is that by taking successively finer partitions of the image of f one can better approximate f.

For the elementary function we are depicting in Figure 5.1, the two approaches appear to be much the same. However, the power of the Lebesgue integral rests in its use of the "size" of the sets A_j on which the approximating function is constant. For step functions, these sets are always intervals, and it is there that the problems arise. By allowing the sets A_j to be quite general, the Lebesgue integral becomes a very powerful tool. However, it does need some buildup, and the first thing to do is remove the quotes from "size."

5.1.4 Notes

Example 5.1.9 comes from [**RB:99**]. **OF:33** connects the Riemann integral and the Jordan measure.

[**PRH:74, DLC:13**]

Exercises

5.1.1 Prove Proposition 5.1.8.

Section 5.2

Measurable sets

The construction of the integral we provide in this chapter proceeds along different lines than does the usual construction of the Riemann integral. In Riemann integration one typically jumps right in with a function and starts constructing step function approximations, etc. However, one could also define the Riemann integral by first defining the Jordan measure as in Section 5.1.1, and then using this as the basis for defining the integral. But the idea is still that one uses step functions as approximations. In the theory for integration that we develop here, a crucial difference is the sort of functions we use to approximate the functions we wish to integrate. The construction of these approximating functions, in turn, rests on some purely set theoretic constructions that play the rôle of the Jordan measure (which, we remind the reader, is not a measure in the general sense we define in this chapter) in Riemann integration. In this section we provide the set theoretic constructions needed to begin this abstract form of integration theory.

Do I need to read this section? If you are reading this chapter, then this is where the technical material begins. If you are only interested in learning about Lebesgue measure, you can get away with knowing the definition of "measurable space" and then proceeding directly to Section 5.4. However, in Section 5.4 we will freely refer to things proved in this section, so as you read Section 5.4 you will eventually end up reading many things in this section anyway. •

5.2.1 Algebras and σ-algebras

The idea we develop in this section and the next is that of a means of measuring the size of a set in a general way. What one first must do is provide a suitable collection of sets whose size one wishes to measure. One's first reaction to this programme might be, "Why not measure the size of *all* subsets?" The answer to this question is not immediately obvious, and we shall say some things about this as we go along. For the moment, the reader should simply trust that the definitions we give have been thought over pretty carefully by lots of pretty smart people, and so are possibly "correct."[2]

5.2.1 Definition (Algebra, σ-algebra, measurable space) For a set X, a subset of subsets $\mathscr{A} \subseteq \mathbf{2}^X$ is an ***algebra***[3] if

(i) $X \in \mathscr{A}$,

(ii) $A \in \mathscr{A}$ implies $X \setminus A \in \mathscr{A}$, and

(iii) $\cup_{j=1}^{k} A_j \in \mathscr{A}$ for any finite family $(A_1, \dots, A_k)$ of subsets,

[2] That being said, *never* stop being a skeptic!

[3] Also sometimes called a ***field***.

and a ***σ-algebra***[4] on X if

(iv) $X \in \mathscr{A}$,

(v) $A \in \mathscr{A}$ implies $X \setminus A \in \mathscr{A}$, and

(vi) $\cup_{j \in \mathbb{Z}_{>0}} A_j \in \mathscr{A}$ for any countable family $(A_j)_{j \in \mathbb{Z}_{>0}}$ of subsets.

A pair $(X, \mathscr{A})$ is called a ***measurable space*** if $\mathscr{A}$ is a σ-algebra on X and elements of $\mathscr{A}$ are called ***$\mathscr{A}$-measurable***. •

We shall mainly be concerned with σ-algebras, although the notion of an algebra is occasionally useful even if one is working with σ-algebras.

5.2.2 Remark (Why are the axioms for a measurable space as they are?) In Remark **??** we attempted to justify why the axioms for a topological space are as they are. For topological spaces this justification is facilitated by the fact that most readers will already know about open subsets of Euclidean space. For readers new to measure theory, it is less easy to justify the axioms of a measurable space. In particular, why is it that we require *countable* unions of measurable subsets to be measurable? Why not finite unions (as with algebras) or arbitrary unions? Why not intersections instead of unions? The reason for this, at its core, is that we wish for the theory we develop to have useful properties with respect to sequential limit operations, and such limit operations have an intrinsic countability in them due to sequences being countable sets. It may be difficult to see just why this is important at this point, but this is the justification. •

Let us give some simple examples of σ-algebras.

5.2.3 Examples (Algebras, σ-algebras)

1. It is clear that the power set $\mathbf{2}^X$ of a set X is a σ-algebra.
2. For a set X, the collection of subsets $\{\emptyset, X\}$ is a σ-algebra.
3. For a set X the collection of subsets

$$\mathscr{A} = \{A \subseteq X \mid A \text{ or } X \setminus A \text{ is countable}\}$$

is a σ-algebra.

4. The collection $\mathscr{J}(\mathbb{R}^n)$ of Jordan measurable subsets of $\mathbb{R}^n$ (see Definition **??**) is an algebra by Proposition 5.1.8 and not a σ-algebra by virtue of Example 5.1.10. •

The following result records some useful properties of σ-algebras.

5.2.4 Proposition (Properties of σ-algebras) *Let $\mathscr{A}$ be a σ-algebra on* X. *The following statements hold:*

(i) $\emptyset \in \mathscr{A}$;

(ii) if $\mathrm{A}_1, \dots, \mathrm{A}_\mathrm{k} \in \mathscr{A}$ *then* $\cup_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{A}_\mathrm{j} \in \mathscr{A}$;

(iii) $\cap_{\mathrm{j} \in \mathbb{Z}_{>0}} \mathrm{A}_\mathrm{j} \in \mathscr{A}$ *for any countable collection* $(\mathrm{A}_\mathrm{j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *of subsets;*

[4] Also sometimes called a ***σ-field***.

(iv) *if* $\mathrm{A}_1, \dots, \mathrm{A}_k \in \mathscr{A}$ *then* $\cap_{j=1}^{k} \mathrm{A}_j \in \mathscr{A}$.

Moreover, condition (vi) in Definition 5.2.1 can be equivalently replaced with condition (iii) above.

Proof (i) Since $X \in \mathscr{A}$ we must have $X \setminus X = \emptyset \in \mathscr{A}$.

(ii) We define a countable collection $(B_j)_{j\in\mathbb{Z}_{>0}}$ of subsets in $\mathscr{A}$ by

$$B_j = \begin{cases} A_j, & j \in \{1, \dots, k\}, \\ \emptyset, & j > k, \end{cases}$$

and the assertion now follows since

$$\cup_{j=1}^{k} A_j = \cup_{j\in\mathbb{Z}_{>0}} B_j \in \mathscr{A}.$$

(iii) This follows from De Morgan's Laws (Proposition 1.1.5):

$$\bigcap_{j\in\mathbb{Z}_{>0}} A_j = X \setminus \Big(\bigcup_{j\in\mathbb{Z}_{>0}} (X \setminus A_j) \Big).$$

Since $X \setminus A_j \in \mathscr{A}$ it follows that $\cup_{j\in\mathbb{Z}_{>0}} (X \setminus A_j) \in \mathscr{A}$ since $\mathscr{A}$ is a σ-algebra. Therefore $X \setminus \big(\cup_{j\in\mathbb{Z}_{>0}} (X \setminus A_j) \big) \in \mathscr{A}$ and so this part of the result follows.

(iv) This follows again from De Morgans's Laws, along with part (ii).

The final assertion of the proposition follows from De Morgans's Laws, as can be gleaned from the arguments used in the proof of part (iii), along with a similar argument, swapping the rôles of union and intersection. ■

The following corollary is now obvious.

5.2.5 Corollary (σ-algebras are algebras) *A σ-algebra $\mathscr{A}$ on a set* X *is also an algebra on* X.

Another construction that is sometimes useful is the restriction of a measurable space $(X, \mathscr{A})$ to a subset $A \subseteq X$. If A is measurable, then there is a natural σ-algebra induced on A.

5.2.6 Proposition (Restriction of a σ-algebra to a measurable subset) *Let* $(\mathrm{X}, \mathscr{A})$ *be a measurable space, let* $\mathrm{A} \in \mathscr{A}$*, and define* $\mathscr{A}_\mathrm{A} \subseteq \mathbf{2}^\mathrm{A}$ *by*

$$\mathscr{A}_\mathrm{A} = \{\mathrm{B} \cap \mathrm{A} \mid \mathrm{B} \in \mathscr{A}\}.$$

Then $(\mathrm{A}, \mathscr{A}_\mathrm{A})$ *is a measurable space.*

Proof We need to show that $\mathscr{A}_A$ is a σ-algebra on A. Clearly $A \in \mathscr{A}_A$ since $A = X \cap A$ and $X \in \mathscr{A}$. Also, since $A \setminus (B \cap A) = (X \setminus B) \cap A$ by Proposition 1.1.5, it follows that $A \setminus (B \cap A) \in \mathscr{A}_A$ for $B \cap A \in \mathscr{A}_A$. Suppose that $(B_j \cap A)_{j\in\mathbb{Z}_{>0}}$ is a countable family of sets in $\mathscr{A}_A$. Since $\cup_{j\in\mathbb{Z}_{>0}} (B_j \cap A) = (\cup_{j\in\mathbb{Z}_{>0}} B_j) \cap A$ by Proposition 1.1.7 it follows that $\cup_{j\in\mathbb{Z}_{>0}} (B_j \cap A) \in \mathscr{A}_A$. ■

5.2.2 Algebras and σ-algebras generated by families of subsets

It is often useful to be able to indirectly define algebras and σ-algebras by knowing that they contain a certain family of subsets. This is entirely analogous to the manner in which one defines a topology by a basis or subbasis; see Section **??**.

Let us begin with the construction of a σ-algebra containing a family of subsets.

5.2.7 Proposition (σ-algebras generated by subsets) *If X is a set and if $\mathscr{S} \subseteq \mathbf{2}^X$ then there exists a unique σ-algebra $\sigma(\mathscr{S})$ with the following properties:*

(i) $\mathscr{S} \subseteq \sigma(\mathscr{S})$;

(ii) if $\mathscr{A}$ is any σ-algebra for which $\mathscr{S} \subseteq \mathscr{A}$ then $\sigma(\mathscr{S}) \subseteq \mathscr{A}$.

Proof We let $\mathscr{P}_{\mathscr{S}}$ be the collection of all σ-algebras with the property that if $\mathscr{A} \in \mathscr{P}_{\mathscr{S}}$ then $\mathscr{S} \subseteq \mathscr{A}$. Note that $\mathscr{P}_{\mathscr{S}}$ is nonempty since $\mathbf{2}^X \subseteq \mathscr{P}_{\mathscr{S}}$. We then define

$$\sigma(\mathscr{S}) = \bigcap \{\mathscr{A} \mid \mathscr{A} \in \mathscr{P}_{\mathscr{S}}\}.$$

If $\sigma(\mathscr{S})$ is a σ-algebra then clearly it satisfies the conditions of the statement of the result. Let us then show that $\sigma(\mathscr{S})$ is a σ-algebra. Since each element of $\mathscr{P}_{\mathscr{S}}$ is a σ-algebra we have $X \in \mathscr{A}$ whenever $\mathscr{A} \in \mathscr{P}_{\mathscr{S}}$. Therefore $X \in \sigma(\mathscr{S})$. If $A \in \sigma(\mathscr{S})$ it follows that $A \in \mathscr{A}$ whenever $\mathscr{A} \in \mathscr{P}_{\mathscr{S}}$. Therefore $X \setminus A \in \mathscr{A}$ whenever $\mathscr{A} \in \mathscr{P}_{\mathscr{S}}$, showing that $X \setminus A \in \sigma(\mathscr{S})$. Finally, if $(A_j)_{j\in\mathbb{Z}_{>0}} \subseteq \sigma(\mathscr{S})$ then $(A_j)_{j\in\mathbb{Z}_{>0}} \subseteq \mathscr{A}$ whenever $\mathscr{A} \in \mathscr{P}_{\mathscr{S}}$. Therefore, $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{A}$ whenever $\mathscr{A} \in \mathscr{P}_{\mathscr{S}}$. Therefore, $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \sigma(\mathscr{S})$. ■

The previous proof applies equally well to algebras. Moreover, it is possible to give a more or less explicit characterisation of the smallest algebra containing a given collection of subsets. This is not possible for σ-algebras, cf. the proof of Theorem 5.2.14.

5.2.8 Proposition (Algebras generated by subsets) *If X is a set and if $\mathscr{S} \subseteq \mathbf{2}^X$ then there exists a unique algebra $\sigma_0(\mathscr{S})$ with the following properties:*

(i) $\mathscr{S} \subseteq \sigma_0(\mathscr{S})$;

(ii) if $\mathscr{A}$ is any algebra for which $\mathscr{S} \subseteq \mathscr{A}$ then $\sigma_0(\mathscr{S}) \subseteq \mathscr{A}$.

Moreover, $\sigma_0(\mathscr{S})$ is the set of finite unions of sets of the form $S_1 \cap \cdots \cap S_k$, where each of the sets $S_1, \dots, S_k$ is either in $\mathscr{S}$ or its complement is in $\mathscr{S}$.

Proof The existence of $\sigma_0(\mathscr{S})$ can be argued just as in the proof of Proposition 5.2.7. To see that $\sigma_0(\mathscr{S})$ admits the explicit stated form, let $\overline{\mathscr{S}}$ be the collection sets of the stated form. We first claim that $\overline{\mathscr{S}}$ is an algebra. To see that $X \in \overline{\mathscr{S}}$, let $S \in \mathscr{S}$ and note that $X \setminus S \in \overline{\mathscr{S}}$. Thus $X = S \cup (X \setminus S) \in \overline{\mathscr{S}}$. If $T \in \overline{\mathscr{S}}$ then we show that $X \setminus T \in \overline{\mathscr{S}}$ as follows. Note that $T = T_1 \cup \cdots \cup T_k$ where, for each $j \in \{1, \dots, k\}$,

$$T_j = \bigcap_{l_j=1}^{m_j} S_{jl_j}, \qquad S_{jl_j} \in \mathscr{S} \text{ or } X \setminus S_{jl_j} \in \mathscr{S},\ l_j \in \{1, \dots, m_j\}.$$

Let us for brevity denote $A = \{1, \dots, m_1\} \times \cdots \times \{1, \dots, m_k\}$. Then, using De Morgan's Laws and Proposition 1.1.7,

$$\begin{aligned} X \setminus T = X \setminus \Big(\bigcup_{j=1}^{k} \Big(\bigcap_{l_j=1}^{m_j} S_{jl_j}\Big)\Big) &= \bigcap_{j=1}^{k} \Big(X \setminus \Big(\bigcap_{l_j=1}^{m_j} S_{jl_j}\Big)\Big) \\ &= \bigcap_{j=1}^{k} \Big(\bigcup_{l_j=1}^{m_j} X \setminus S_{jl_j}\Big) = \bigcup_{(l_1,\dots,l_k)\in A} \Big(\bigcup_{j=1}^{k} X \setminus S_{jl_j}\Big), \end{aligned}$$

which then gives $X \setminus T \in \overline{\mathscr{S}}$. It is obvious that finite unions of sets from $\overline{\mathscr{S}}$ are in $\overline{\mathscr{S}}$, which shows that $\overline{\mathscr{S}}$ is an algebra, as desired. Moreover, it is clear that $\mathscr{S} \subseteq \overline{\mathscr{S}}$.

Now suppose that $\mathscr{A}$ is an algebra for which $\mathscr{S} \subseteq \mathscr{A}$. Since $\mathscr{A}$ is an algebra this implies that $X \setminus S \in \mathscr{A}$ for $S \in \mathscr{S}$ and, by Exercise 5.2.1, that $S_1 \cap \cdots \cap S_k \in \mathscr{A}$ for every collection $S_1, \dots, S_k$ for which either $S_j \in \mathscr{S}$ or $X \setminus S_j \in \mathscr{S}$ for each $j \in \{1, \dots, k\}$. Thus $\overline{\mathscr{S}} \subseteq \mathscr{A}$ and so $\overline{\mathscr{S}} = \sigma_0(\mathscr{S})$, as desired. ■

This gives the following result as a special case.

5.2.9 Corollary (The algebra generated by a finite collection of sets) *Let* X *be a set and let* $\mathrm{S}_1, \dots, \mathrm{S}_k \subseteq \mathrm{X}$ *be a finite family of subsets. Then* $\sigma_0(\mathrm{S}_1, \dots, \mathrm{S}_k)$ *is the collection of finite unions of sets of the form* $\mathrm{T}_1 \cap \cdots \cap \mathrm{T}_m$ *where, for each* $\mathrm{j} \in \{1, \dots, \mathrm{m}\}$*, either* $\mathrm{T}_j \in \{\mathrm{S}_1, \dots, \mathrm{S}_k\}$ *or* $\mathrm{X} \setminus \mathrm{T}_j \in \{\mathrm{S}_1, \dots, \mathrm{S}_k\}$.

The point is that you can specify any collection of subsets and define an algebra or σ-algebra associated with this collection in a natural way, i.e., by demanding that the conditions of an algebra or a σ-algebra hold. The preceding results makes sense of the next definition.

5.2.10 Definition (Algebras and σ-algebras generated by subsets) If X is a set and $\mathscr{S} \subseteq \mathbf{2}^X$, the algebra $\sigma_0(\mathscr{S})$ (resp. σ-algebra $\sigma(\mathscr{S})$) of Proposition 5.2.8 (resp. Proposition 5.2.7) is the ***algebra generated by*** $\mathscr{S}$ (resp. ***σ-algebra generated by*** $\mathscr{S}$). •

We now provide an alternative description of the σ-algebra generated by a collection of subsets. This description relies on the following concept.

5.2.11 Definition (Monotone class) For a set X, a ***monotone class*** on X is a collection $\mathscr{M} \subseteq \mathbf{2}^X$ of subsets of X with the following properties:

(i) $\cup_{j \in \mathbb{Z}} A_j \in \mathscr{M}$ for every family $(A_j)_{j \in \mathbb{Z}_{>0}}$ of subsets from $\mathscr{M}$ such that $A_j \subseteq A_{j+1}$ for every $j \in \mathbb{Z}_{>0}$;

(ii) $\cap_{j \in \mathbb{Z}_{>0}} A_j \in \mathscr{M}$ for every family $(A_j)_{j \in \mathbb{Z}_{>0}}$ of subsets from $\mathscr{M}$ such that $A_j \supseteq A_{j+1}$ for every $j \in \mathbb{Z}_{>0}$. •

Let us illustrate how the conditions of a monotone class can be used to relate algebras and σ-algebras.

5.2.12 Proposition (Algebras that are σ-algebras) *Let* X *be a set and let* $\mathscr{A}$ *be an algebra. If either*

(i) $\cup_{j \in \mathbb{Z}_{>0}} \mathrm{A}_j \in \mathscr{A}$ *for every family* $(\mathrm{A}_j)_{j \in \mathbb{Z}_{>0}}$ *of subsets from* $\mathscr{A}$ *for which* $\mathrm{A}_j \subseteq \mathrm{A}_{j+1}$, $\mathrm{j} \in \mathbb{Z}_{>0}$*, or*

(ii) $\cap_{j \in \mathbb{Z}_{>0}} \mathrm{A}_j \in \mathscr{A}$ *for every family* $(\mathrm{A}_j)_{j \in \mathbb{Z}_{>0}}$ *of subsets from* $\mathscr{A}$ *for which* $\mathrm{A}_j \supseteq \mathrm{A}_{j+1}$, $\mathrm{j} \in \mathbb{Z}_{>0}$,

then $\mathscr{A}$ *is a* σ*-algebra.*

Proof We clearly have $X \in \mathscr{A}$ and $X \setminus A \in \mathscr{A}$ for $A \in \mathscr{A}$.

Now suppose that the first of the two conditions in the proposition holds and let $(A_j)_{j \in \mathbb{Z}_{>0}}$ be a countable collection of subsets from $\mathscr{A}$. For $k \in \mathbb{Z}_{>0}$ define $B_k \in \cup_{j=1}^k A_j$. Since $\mathscr{A}$ is an algebra, $B_k \in \mathscr{A}$ for $k \in \mathbb{Z}_{>0}$. Moreover, we clearly have $B_k \subseteq B_{k+1}$ for each $k \in \mathbb{Z}_{>0}$ and $\cup_{j\in\mathbb{Z}_{>0}} A_j = \cup_{k\in\mathbb{Z}_{>0}} B_k$. Therefore, by assumption, $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{A}$, and so $\mathscr{A}$ is a σ-algebra.

Finally suppose that the second of the two conditions in the proposition holds and let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a countable collection of subsets from $\mathscr{A}$. Define $B_k = X \setminus \cup_{j=1}^{k} A_j$. Since $\mathscr{A}$ is an algebra we have $B_k \in \mathscr{A}$ for $k \in \mathbb{Z}_{>0}$. We also have $B_k \supseteq B_{k+1}$ for each $k \in \mathbb{Z}_{>0}$ and $\cap_{k=1}^{\infty} = X \setminus \cup_{j\in\mathbb{Z}_{>0}} A_j$. Thus $X \setminus \cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{A}$, and so $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{A}$ since $\mathscr{A}$ is an algebra. Thus $\mathscr{A}$ is a σ-algebra. ■

Next we state our alternative characterisation of the σ-algebra generated by an algebra of subsets. It is perhaps not immediately apparent why the result is useful, but we shall use it in our discussion of product measures in Section 5.8.1.

5.2.13 Theorem (Monotone Class Theorem) *Let* X *be a set and let* $\mathscr{S} \subseteq \mathbf{2}^{\mathsf{X}}$. *Then there exists a unique monotone class* $\mathrm{m}(\mathscr{S})$ *on* X *such that*

(i) $\mathscr{S} \subseteq \mathrm{m}(\mathscr{S})$ *and*

(ii) if $\mathscr{M}$ *is any monotone class on* X *for which* $\mathscr{S} \subseteq \mathscr{M}$ *then* $\mathrm{m}(\mathscr{S}) \subseteq \mathscr{M}$.

Moreover, if $\mathscr{S}$ *is an algebra then* $\mathrm{m}(\mathscr{S}) = \sigma(\mathscr{S})$.

Proof We let $\mathscr{P}_{\mathscr{S}}$ be the collection of monotone classes with the property that if $\mathscr{M} \in \mathscr{P}_{\mathscr{S}}$ then $\mathscr{S} \subseteq \mathscr{M}$. Since $X \in \mathscr{P}_{\mathscr{S}}$ it follows that $\mathscr{P}_{\mathscr{S}}$ is not empty. We define

$$m(\mathscr{S}) = \bigcap \{\mathscr{M} \mid \mathscr{M} \in \mathscr{P}_{\mathscr{S}}\}.$$

It is clear that $\mathscr{S} \subseteq m(\mathscr{S})$. Moreover, it is also clear that if $\mathscr{M}$ is a monotone class containing $\mathscr{S}$ then $m(\mathscr{S}) \subseteq \mathscr{M}$. It remains to show that $m(\mathscr{S})$ is a monotone class. Let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a family of subsets from $m(\mathscr{S})$ such that $A_j \subseteq A_{j+1}$ for $j \in \mathbb{Z}_{>0}$. Since $A_j \in \mathscr{M}$ for each $j \in \mathbb{Z}_{>0}$ and $\mathscr{M} \in \mathscr{P}_{\mathscr{S}}$ it follows that $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{M}$ for every $\mathscr{M} \in \mathscr{P}_{\mathscr{S}}$. Thus $\cup_{j\in\mathbb{Z}_{>0}} A_j \in m(\mathscr{S})$. Similarly, let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a family of subsets from $m(\mathscr{S})$ for which $A_j \supseteq A_{j+1}$ for $j \in \mathbb{Z}_{>0}$. Since $A_j \in \mathscr{M}$ for every $j \in \mathbb{Z}_{>0}$ and $\mathscr{M} \in \mathscr{P}_{\mathscr{S}}$ it follows that $\cap_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{M}$ for every $\mathscr{M} \in \mathscr{P}_{\mathscr{S}}$. Thus $\cap_{j\in\mathbb{Z}_{>0}} A_j \in m(\mathscr{S})$, showing that $m(\mathscr{S})$ is indeed a monotone class.

Now let us prove the final assertion of the theorem, supposing that $\mathscr{S}$ is an algebra. We claim that $m(\mathscr{S})$ is an algebra. Indeed, let $S \in \mathscr{S}$ and define

$$\mathscr{M}_S = \{A \in m(\mathscr{S}) \mid S \cap A, S \cap (X \setminus A), (X \setminus S) \cap A \in m(\mathscr{S})\}.$$

We claim that $\mathscr{M}_S$ is a monotone class. Indeed, let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a family of subsets from $\mathscr{M}_S$ such that $A_j \subseteq A_{j+1}$ for $j \in \mathbb{Z}_{>0}$. Thus

$$S \cap A_j, S \cap (X \setminus A_j), (X \setminus S) \cap A_j \in m(\mathscr{S}), \qquad j \in \mathbb{Z}_{>0}.$$

Then, using Propositions 1.1.5 and 1.1.7,

$$\begin{aligned}
&S \cap \left(\cup_{j\in\mathbb{Z}_{>0}} A_j\right) = \cup_{j\in\mathbb{Z}_{>0}} (S \cap A_j),\\
&S \cap \left(X \setminus \left(\cup_{j\in\mathbb{Z}_{>0}} A_j\right)\right) = S \cap \left(\cap_{j\in\mathbb{Z}_{>0}} X \setminus A_j\right) = \cap_{j\in\mathbb{Z}_{>0}} S \cap (X \setminus A_j),\\
&(X \setminus S) \cap \left(\cup_{j\in\mathbb{Z}_{>0}} A_j\right) = \cup_{j\in\mathbb{Z}_{>0}} (X \setminus S) \cap A_j.
\end{aligned}$$

Since

$$S \cap A_j \subseteq S \cap A_{j+1}, \quad S \cap (X \setminus A_j) \supseteq S \cap (X \setminus A_{j+1}), \quad (X \setminus S) \cap A_j \subseteq (X \setminus S) \cap A_{j+1},$$

for $j \in \mathbb{Z}_{>0}$, we conclude that

$$S \cap \left(\cup_{j\in\mathbb{Z}_{>0}} A_j\right), S \cap \left(X \setminus \left(\cup_{j\in\mathbb{Z}_{>0}} A_j\right)\right), (X \setminus S) \cap \left(\cup_{j\in\mathbb{Z}_{>0}} A_j\right) \in m(\mathscr{S}),$$

and so $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{M}_S$. A similarly styled argument gives $\cap_{j\in\mathbb{Z}_{>0}} \in \mathscr{M}_S$ for a countable family $(A_j)_{j\in\mathbb{Z}_{>0}}$ of subsets from $\mathscr{M}_S$ satisfying $A_j \supseteq A_{j+1}$, $j \in \mathbb{Z}_{>0}$. Thus $\mathscr{M}_S$ is indeed a monotone class.

We claim that $\mathscr{M}_S = m(\mathscr{S})$. To see this we first claim that $\mathscr{S} \subseteq \mathscr{M}_S$. Indeed, if $A \in \mathscr{S}$ then

$$S \cap A, S \cap (X \setminus A), (X \setminus S) \cap A \in \mathscr{S} \subseteq m(\mathscr{S})$$

since $\mathscr{S}$ is a field. Thus $\mathscr{M}_S$ is a monotone class containing $\mathscr{S}$ and so $m(\mathscr{S}) \subseteq \mathscr{M}_S$. Since $\mathscr{M}_S \subseteq m(\mathscr{S})$ by definition, we conclude that $\mathscr{M}_S = m(\mathscr{S})$. Note that $S \in \mathscr{S}$ is arbitrary in this construction.

Next we claim that $\mathscr{M}_S$, and so $m(\mathscr{S})$, is an algebra. First of all, since $X \in \mathscr{S}$ by virtue of $\mathscr{S}$ being an algebra, we have

$$X \in \mathscr{S} \subseteq m(\mathscr{S}) = \mathscr{M}_S.$$

Also, if $A \in \mathscr{M}_S$ we have

$$A \in \mathscr{M}_S \implies A \in \mathscr{M}_X \implies X \cap (X \setminus A) \in m(\mathscr{S}) \implies X \setminus A \in m(\mathscr{S}) = \mathscr{M}_S.$$

Also, let $A, B \in \mathscr{M}_S$. Then

$$A, B \in \mathscr{M}_S \implies A, B \in \mathscr{M}_A \implies B \cap A \in m(\mathscr{S}) = \mathscr{M}_S.$$

Thus the intersection of sets from $\mathscr{M}_S$ lies in $\mathscr{M}_S$. This means that if $A, B \in \mathscr{M}_S$ then

$$X \setminus A, X \setminus B \in \mathscr{M}_S \implies (X \setminus A) \cap (X \setminus B) = X \setminus (A \cup B) \in \mathscr{M}_S,$$

implying that $A \cup B \in \mathscr{M}_S$. Thus pairwise unions of sets from $\mathscr{M}_X$ are in $\mathscr{M}_S$. An elementary induction then gives $\cap_{j=1}^k A_j \in \mathscr{M}_S$ for every family of subsets $(A_1, \dots, A_j)$ from $\mathscr{M}_S$. This shows that $\mathscr{M}_S = m(\mathscr{S})$ is an algebra.

Since $\mathscr{M}_S$ is a monotone class it is a σ-algebra by Proposition 5.2.12. Thus $\sigma(\mathscr{S}) \subseteq \mathscr{M}_S = m(\mathscr{S})$. Moreover, $\sigma(\mathscr{S})$ is a monotone class by the properties of a σ-algebra and by Proposition 5.2.4. Since $\mathscr{S} \subseteq \sigma(\mathscr{S})$ we conclude from Proposition 5.2.12 that $m(\mathscr{S}) \subseteq \sigma(\mathscr{S})$, giving $m(\mathscr{S}) = \sigma(\mathscr{S})$, as desired. ■

The following "fun fact" about the σ-algebra generated by a collection of subsets is useful to understand how big this σ-algebra is. We will use this result in Proposition 5.4.13 to compare the cardinalities of Borel and Lebesgue measurable sets. Recall that $\aleph_0 = \operatorname{card}(\mathbb{Z}_{\geq 0})$.

5.2.14 Theorem (Cardinality of the σ-algebra generated by a collection of subsets) *Let* X *be a set and let* $\mathscr{S} \subseteq \mathbf{2}^X$ *be such that* $\emptyset \in \mathscr{S}$ *and that* $\operatorname{card}(\mathscr{S}) \geq 2$. *Then* $\operatorname{card}(\sigma(\mathscr{S})) \leq \operatorname{card}(\mathscr{S})^{\aleph_0}$.

Proof Let $\aleph_1$ be the smallest uncountable cardinal number (the cardinal number that the Continuum Hypothesis asserts is equal to $\operatorname{card}(\mathbb{R})$). Define $\mathscr{S}_0 = \mathscr{S}$. For a cardinal number $c < \aleph_1$ we shall use Transfinite Induction (Theorem **??**) to define $\mathscr{S}_c$ as follows. Suppose that $\mathscr{S}_{c'}$ has been defined for a cardinal number c' such that $0 < c' < c$. Then define $\mathscr{S}_c$ to be the collection of sets of the form $\cup_{j\in\mathbb{Z}_{>0}}A_j$ where either A_j or $X \setminus A_j$ is an element of the family $\cup_{0\le c'<c}\mathscr{S}_{c'}$ of subsets of X. We claim that $\cup_{0\le c<\aleph_1}\mathscr{S}_c = \sigma(\mathscr{S})$.

We first prove by Transfinite Induction that $\cup_{0\le c<\aleph_1}\mathscr{S}_c \subseteq \sigma(\mathscr{S})$. Clearly $\mathscr{S}_0 \subseteq \sigma(\mathscr{S})$. Suppose that $\mathscr{S}_{c'} \subseteq \sigma(\mathscr{S})$ for $0 \le c' < c < \aleph_1$. Then let $\cup_{j\in\mathbb{Z}_{>0}}A_j \in \mathscr{S}_c$ for set A_j such that either A_j or $X \setminus A_j$ are in the family $\cup_{0\le c'<c}\mathscr{S}_{c'}$ of subsets of X. It follows from the induction hypothesis that $A_j, X \setminus A_j \in \sigma(\mathscr{S})$. Thus $\cup_{j\in\mathbb{Z}_{>0}}A_j \in \sigma(\mathscr{S})$ since a σ-algebra is closed under countable unions. Therefore, $\mathscr{S}_c \in \sigma(\mathscr{S})$ and so we conclude from Transfinite Induction that $\cup_{0\le c<\aleph_1}\mathscr{S}_c \subseteq \sigma(\mathscr{S})$.

To prove that $\cup_{0\le c<\aleph_1}\mathscr{S}_c = \sigma(\mathscr{S})$ it now suffices to show that $\cup_{0\le c<\aleph_1}\mathscr{S}_c$ is a σ-algebra since it contains $\mathscr{S}$ and since $\sigma(\mathscr{S})$ is the smallest σ-algebra containing $\mathscr{S}$. Since $\emptyset \in \mathscr{S}$ we have

$$X = (X \setminus \emptyset) \cup \emptyset \cup \emptyset \cdots \in \mathscr{S}_1,$$

and so $X \in \cup_{0\le c<\aleph_1}\mathscr{S}_c$. Now suppose that $A \in \cup_{0\le c<\aleph_1}\mathscr{S}_c$ so that $A \in \mathscr{S}_{c_0}$ for some c_0 satisfying $0 \le c_0 \le \aleph_1$. For $c_1 > c_0$ it then holds that

$$X \setminus A = (X \setminus A) \cup (X \setminus A) \cup \cdots \in \mathscr{S}_{c_1},$$

and so $(X \setminus A) \in \cup_{0\le c<\aleph_1}\mathscr{S}_c$. Finally, let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a countable family of subsets from $\cup_{0\le c<\aleph_1}\mathscr{S}_c$. For $j \in \mathbb{Z}_{>0}$ let c_j be a cardinal number satisfying $0 \le c_j < \aleph_1$ and $A_j \in \mathscr{S}_{c_j}$. Since $\aleph_1$ is uncountable it cannot be a countable union of countable sets (by Proposition **??**) and since each of the cardinal numbers c_j, $j \in \mathbb{Z}_{>0}$, are countable, it follows that there exists a cardinal number c_∞ such that $0 \le c_\infty < \aleph_1$ and such that $c_j < c_\infty$. Then $\cup_{j\in\mathbb{Z}_{>0}}A_j \in \mathscr{S}_{c_\infty} \subseteq \cup_{0\le c<\aleph_1}\mathscr{S}_c$, completing the proof that $\cup_{0\le c<\aleph_1}\mathscr{S}_c = \sigma(\mathscr{S})$.

We now prove by Transfinite Induction that $\operatorname{card}(\mathscr{S}_c) \le \operatorname{card}(\mathscr{S})^{\aleph_0}$ for every cardinal number c satisfying $0 \le c \le \aleph_1$. Certainly $\operatorname{card}(\mathscr{S}_0) \le \operatorname{card}(\mathscr{S})^{\aleph_0}$. Now suppose that c is a cardinal number satisfying $0 \le c < \aleph_1$ and suppose that $\operatorname{card}(\mathscr{S}_{c'}) \le \operatorname{card}(\mathscr{S})^{\aleph_0}$ for cardinals c' satisfying $0 \le c' < c$. Since c is countable it follows that

$$\operatorname{card}(\cup_{0\le c'<c}\mathscr{S}_{c'}) \le \aleph_0 \operatorname{card}(\mathscr{S})^{\aleph_0} = \operatorname{card}(\mathscr{S})^{\aleph_0}$$

by Theorem **??**, Exercises **??** and **??**, and since $\operatorname{card}(\mathscr{S}) \ge 2$. Now, considering the definition of $\mathscr{S}_c$ we see that

$$\operatorname{card}(\mathscr{S}_c) = 2\operatorname{card}(\cup_{0\le c'<c}\mathscr{S}_{c'}) \le \operatorname{card}(\mathscr{S})^{\aleph_0},$$

as claimed.

From this we deduce that

$$\operatorname{card}(\sigma(\mathscr{S})) = \operatorname{card}(\cup_{0\le c<\aleph_1}\mathscr{S}_c) \le \operatorname{card}(\mathscr{S})^{\aleph_0}\aleph_1 = \operatorname{card}(\mathscr{S})^{\aleph_0},$$

using Theorem **??** and the fact that $\operatorname{card}(\mathscr{S})^{\aleph_0} \ge \aleph_1$ since $\operatorname{card}(\mathscr{S}) \ge 2$ and using Exercises **??** and **??**. ■

5.2.3 Products of measurable spaces

The development of measure theory on products is a little more challenging than, say, the development of topology on products. In this section we introduce the basic tool for studying measure theory for products by considering the products of sets equipped with algebras or σ-algebras of subsets.

We begin by considering products of sets equipped with algebras of subsets.

5.2.15 Definition (Measurable rectangles) For sets $X_1, \dots, X_k$ with algebras $\mathscr{A}_j \subseteq \mathbf{2}^{X_j}$, $j \in \{1, \dots, k\}$, a ***measurable rectangle*** is a subset

$$A_1 \times \cdots \times A_k \subseteq X_1 \times \cdots \times X_k$$

where $A_j \in \mathscr{A}_j$, $j \in \{1, \dots, k\}$. The set of measurable rectangles is denoted by $\mathscr{A}_1 \times \cdots \times \mathscr{A}_k$. •

By Corollary 5.2.5 the preceding definition can be applied to the case when each of the collections of subsets $\mathscr{A}_1, \dots, \mathscr{A}_k$ is a σ-algebra.

The following property of the set of measurable rectangles is then useful.

5.2.16 Proposition (Finite unions of measurable rectangles form an algebra) *For sets $X_1, \dots, X_k$ with algebras $\mathscr{A}_j \subseteq \mathbf{2}^{X_j}$, $j \in \{1, \dots, k\}$, the set of finite unions of sets from $\mathscr{A}_1 \times \cdots \times \mathscr{A}_k$ is an algebra on $X_1 \times \cdots \times X_k$, and is necessarily the algebra $\sigma_0(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k)$.*

Proof Clearly $X_1 \times \cdots \times X_k$ is a measurable rectangle. Next, for measurable rectangles $A_1 \times \cdots \times A_k$ and $B_1 \times \cdots \times B_k$ we have

$$(A_1 \times \cdots \times A_k) \cap (B_1 \times \cdots \times B_k) = (A_1 \cap B_1) \times \cdots \times (A_k \cap B_k).$$

This shows that the intersection of two measurable rectangles is a measurable rectangle. From Proposition 1.1.4 we can then conclude that the intersection of two finite unions of measurable rectangles is a finite union of measurable rectangles. Next let $A_1 \times \cdots \times A_k$ be a measurable rectangle and note that

$$(X_1 \times \cdots \times X_k) \setminus (A_1 \times \cdots \times A_k)$$

is the union of sets of the form $B_1 \times \cdots \times B_k$ where $B_j \in \{A_j, X_j \setminus A_j\}$ and where at least one of the sets B_j is not equal to A_j. That is to say, the complement of a measurable rectangle is a finite union of measurable rectangles. By De Morgan's Laws we then conclude that the complement of a finite union of measurable rectangles is a finite union of measurable rectangles. By Exercise 5.2.1 this proves that the set of finite unions of measurable rectangles is an algebra. Moreover, if $\mathscr{A}$ is any σ-algebra containing $\mathscr{A}_1 \times \cdots \times \mathscr{A}_k$ then $\mathscr{A}$ must necessarily contain finite unions of measurable rectangles. Thus $\mathscr{A}$ is contained in the set of finite unions of measurable rectangles. By Proposition 5.2.8 this means that the algebra of finite unions of measurable rectangles is the algebra generated by $\mathscr{A}_1 \times \cdots \times \mathscr{A}_k$. ■

The principal object of interest to us will be the σ-algebra generated by the measurable rectangles. The following result gives a characterisation of this σ-algebra.

5.2.17 Proposition (The σ-algebra generated by the algebra of measurable rectangles) *For sets $X_1, \dots, X_k$ with algebras $\mathscr{A}_j \subseteq \mathbf{2}^{X_j}$, $j \in \{1, \dots, k\}$, we have*

$$\sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k) = \sigma(\sigma(\mathscr{A}_1) \times \cdots \times \sigma(\mathscr{A}_k)).$$

Proof Clearly we have

$$\sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k) \subseteq \sigma(\sigma(\mathscr{A}_1) \times \cdots \times \sigma(\mathscr{A}_k)).$$

To prove the opposite inclusion it suffices to show that

$$\sigma(\mathscr{A}_1) \times \cdots \times \sigma(\mathscr{A}_k) \subseteq \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k)$$

since this will imply that the σ-algebra of the left-hand side is contained in the right-hand side. We prove the preceding inclusion by induction on k. For $k = 1$ the assertion is trivial. So suppose that for $k = m$ we have

$$\sigma(\mathscr{A}_1) \times \cdots \times \sigma(\mathscr{A}_m) \subseteq \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_m),$$

and suppose that we have a set X_{m+1} with an algebra $\mathscr{A}_{m+1}$. Fix $A_j \in \sigma(\mathscr{A}_j)$, $j \in \{1, \dots, m\}$, and define

$$\sigma'(\mathscr{A}_{m+1}) = \{A \in \sigma(\mathscr{A}_{m+1}) \mid A_1 \times \cdots \times A_m \times A \in \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_{m+1})\}.$$

We claim that $\sigma'(\mathscr{A}_{m+1})$ is a σ-algebra on X_{m+1}. Certainly $X_{m+1} \in \sigma'(\mathscr{A}_{m+1})$ since

$$A_1 \times \cdots \times A_m \times X_{m+1} \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_m \times \mathscr{A}_{m+1} \subseteq \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_{m+1}).$$

Let $A \in \sigma'(\mathscr{A}_{m+1})$. Then we note that

$$A_1 \times \cdots \times A_m \times (X_{m+1} \setminus A) = (A_1 \times \cdots \times A_m \times X_{m+1}) \setminus (A_1 \times \cdots \times A_m \times A).$$

By assumption,

$$A_1 \times \cdots \times A_m \times A \in \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_m \times \mathscr{A}_{m+1})$$

from which we conclude that

$$(A_1 \times \cdots \times A_m \times X_{m+1}) \setminus (A_1 \times \cdots \times A_m \times A) \in \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_m \times \mathscr{A}_{m+1}).$$

Thus $X_{m+1} \setminus A \in \sigma'(sA_{m+1})$. Finally, if $(B_j)_{j\in\mathbb{Z}_{>0}}$ is a countable family of subsets from $\sigma'(\mathscr{A}_{m+1})$ we have

$$A_1 \times \cdots \times A_m \times \Big(\bigcup_{j\in\mathbb{Z}_{>0}} B_j\Big) = \bigcup_{j\in\mathbb{Z}_{>0}} A_1 \times \cdots \times A_m \times B_j \in \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_m \times \mathscr{A}_{m+1}).$$

Thus $\cup_{j\in\mathbb{Z}_{>0}} B_j \in \sigma'(\mathscr{A}_{m+1})$, showing that $\sigma'(\mathscr{A}_{m+1})$ is indeed a σ-algebra. Since $\mathscr{A}_{k+1} \subseteq \sigma'(\mathscr{A}_{m+1})$ and since $\sigma'(\mathscr{A}_{m+1}) \subseteq \sigma(\mathscr{A}_{m+1})$, we conclude that $\sigma(\mathscr{A}_{m+1}) = \sigma(\mathscr{A}_{m+1})$. This shows that

$$\sigma(\mathscr{A}_1) \times \cdots \times \sigma(\mathscr{A}_m) \times \sigma(\mathscr{A}_{m+1}) \subseteq \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_m \times \mathscr{A}_{m+1}),$$

as desired. ■

The following property of the product of σ-algebras is useful.

5.2.18 Proposition (Intersections of measurable sets with factors in products are measurable) *Let* $(\mathrm{X}_\mathrm{j}, \mathscr{A}_\mathrm{j})$*,* $\mathrm{j} \in \{1, \dots, \mathrm{k}\}$*, be measurable spaces. For* $\mathrm{A} \in \sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_\mathrm{k})$*, for* $\mathrm{j} \in \{1, \dots, \mathrm{k}\}$*, and for* $\mathrm{x}_\mathrm{j} \in \mathrm{X}_\mathrm{j}$ *define*

$$\mathrm{A}_{\mathrm{x}_\mathrm{j}} = \{(\mathrm{x}_1, \dots, \mathrm{x}_{\mathrm{j}-1}, \mathrm{x}_{\mathrm{j}+1}, \dots, \mathrm{x}_\mathrm{k}) \in \mathrm{X}_1 \times \dots \times \mathrm{X}_{\mathrm{j}-1} \times \mathrm{X}_{\mathrm{j}+1} \times \dots \times \mathrm{X}_\mathrm{k} \mid (\mathrm{x}_1, \dots, \mathrm{x}_{\mathrm{j}-1}, \mathrm{x}_\mathrm{j}, \mathrm{x}_{\mathrm{j}+1}, \dots, \mathrm{x}_\mathrm{k}) \in \mathrm{A}\}.$$

Then $\mathrm{A}_{\mathrm{x}_\mathrm{j}} \in \sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_{\mathrm{j}-1} \times \mathscr{A}_{\mathrm{j}+1} \times \dots \times \mathscr{A}_\mathrm{k})$*.*

Proof Let $\mathscr{F}_{x_j}$ be the subsets $A \subseteq X_1 \times \dots \times X_k$ with the property that $A_{x_j} \in \sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_{j-1} \times \mathscr{A}_{j+1} \times \dots \times \mathscr{A}_k)$. We claim that if $B_j \in X_j$, $j \in \{1, \dots, k\}$, then $B_1 \times \dots \times B_k \in \mathscr{F}_{x_j}$. Indeed, we have $A_{x_j} = B_1 \times \dots \times B_{j-1} \times B_{j+1} \times \dots \times B_k$ if $x_j \in B_j$ and $A_{x_j} = \emptyset$ otherwise. We also claim that $\mathscr{F}_{x_j}$ is a σ-algebra. We have just shown that $X_1 \times \dots \times X_k \in \mathscr{F}_{x_j}$. If $A \in \mathscr{F}_{x_j}$ and $A_l \in \mathscr{F}_{x_j}$, $l \in \mathbb{Z}_{>0}$, then we have the easily verified identities

$$((X_1 \times \dots \times X_k) \setminus A)_{x_j} = (X_1 \times \dots \times X_k) \setminus A_{x_j}$$

and

$$\left(\cup_{l \in \mathbb{Z}_{>0}} A_l\right)_{x_j} = \cup_{j \in \mathbb{Z}_{>0}} (A_l)_{x_j},$$

which shows that $\mathscr{F}_{x_j}$ is indeed a σ-algebra. Since it contains the measurable rectangles we must have

$$\sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_k) \subseteq \mathscr{F}_{x_j}.$$

It, therefore, immediately follows that $A_{x_j} \in \sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_{j-1} \times \mathscr{A}_{j+1} \times \dots \times \mathscr{A}_k)$ whenever $A \in \sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_k$, as desired. ■

Exercises

5.2.1 Let X be a set and let $\mathscr{A}$ be an algebra on X.

(a) Prove the following:

(i) $\emptyset \in \mathscr{A}$;

(ii) if $A_1, \dots, A_k \in \mathscr{A}$ then $\cap_{j=1}^k A_j \in \mathscr{A}$.

(b) Show that condition (iii) in Definition 5.2.1 can be equivalently replaced with condition (ii) above.

5.2.2 Let X be an infinite set. Indicate which of the following collections of subsets are algebras, σ-algebras, or neither:

(a) the collection of finite subsets X;

(b) the collection of subsets A for which $X \setminus A$ is finite;

(c) the collection of countable subsets X;

(d) the collection of subsets A for which $X \setminus A$ is countable.

5.2.3 Answer the following questions.

(a) Is the collection of open subsets of $\mathbb{R}$ an algebra or a σ-algebra?

(b) Is the collection of closed subsets of $\mathbb{R}$ an algebra or a σ-algebra?

5.2.4 Let X be a set and let $\mathscr{S} \subseteq \mathbf{2}^X$. Show that if

$$\mathscr{S}' = \left\{ \bigcup_{j \in \mathbb{Z}_{>0}} A_j \;\middle|\; A_j \in \mathscr{S},\ j \in \mathbb{Z}_{>0} \right\}$$

then the σ-algebras $\sigma(\mathscr{S})$ and $\sigma(\mathscr{S}')$ are generated by $\mathscr{S}$ and $\mathscr{S}'$ agree.

5.2.5 Let X and Y be disjoint sets and let $\mathscr{A}$ and $\mathscr{B}$ be σ-algebras on X and Y, respectively. Let

$$\mathscr{A} \cup \mathscr{B} = \{A \cup B \in \mathbf{2}^{X \cup Y} \mid A \in \mathscr{A},\ B \in \mathscr{B}\}.$$

Show that $\mathscr{A} \cup \mathscr{B}$ is a σ-algebra on $X \cup Y$.

Section 5.3

Measures

The nomenclature "measurable space" from the preceding section makes one think that one ought to be able to measure things in it. This is done with the concept of a measure that we now introduce, and which serves to provide a general framework for talking about the "size" of a subset. The notion of what we shall below call an "outer measure" is perhaps the most intuitive notion of size one can utilise. It has the great advantage of being able to be applied to measure the size of *all* subsets. However, and surprisingly, outer measure has an important defect, namely that it does not have the seemingly natural property of "countable-additivity." The way one gets around this is by restricting outer measure to a collection of subsets where this property of countable-additivity *does* hold. This leads to a natural σ-algebra. At the high level of abstraction in this section, it is not easy to see the justification for the definitions of outer measure and measure. This justification will only become clear in Section 5.4 where there is a fairly intuitive definition of outer measure on $\mathbb{R}$, but that natural outer measure is actually not a measure.

Do I need to read this section? In order to appreciate the framework in which the Lebesgue measure is developed in Sections 5.4 and 5.5, one should understand the notions of measure and outer measure. •

5.3.1 Functions on families of subsets

Before getting to the more specific definitions that we shall mainly use, it is useful to provide some terminology that helps to organise these definitions.

5.3.1 Definition (Properties of functions on subsets) For a set X and a collection $\mathscr{S} \subseteq \mathbf{2}^X$ of subsets of X, a map $\mu\colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$ is:

(i) ***monotonic*** if $\mu(S) \leq \mu(T)$ for subsets $S, T \in \mathscr{S}$ such that $S \subseteq T$;

(ii) ***finitely-subadditive*** if $\mu\Big(\bigcup_{j=1}^{k} S_j\Big) \leq \sum_{j=1}^{k} \mu(S_j)$ for every finite family $(S_1, \dots, S_k)$ of sets from $\mathscr{S}$ whose union is also in $\mathscr{S}$;

(iii) ***countably-subadditive*** if $\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} S_j\Big) \leq \sum_{j=1}^{\infty} \mu(S_j)$ for every countable family $(S_j)_{j\in\mathbb{Z}_{>0}}$ of sets from $\mathscr{S}$ whose union is also in $\mathscr{S}$;

(iv) ***monotonically increasing*** if, for every countable family of subsets $(S_j)_{j\in\mathbb{Z}_{>0}}$ from $\mathscr{S}$ for which $S_j \subseteq S_{j+1}$, $j \in \mathbb{Z}_{>0}$, and whose union is in $\mathscr{S}$, $\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} S_j\Big) = \lim_{j\to\infty} \mu(S_j)$;

(v) ***monotonically decreasing*** if, for every countable family of subsets $(S_j)_{j\in\mathbb{Z}_{>0}}$ from $\mathscr{S}$ for which $S_j \supseteq S_{j+1}$, $j \in \mathbb{Z}_{>0}$, for which $\mu(S_k) < \infty$ for some $k \in \mathbb{Z}_{>0}$, and whose intersection is in $\mathscr{S}$, $\mu\Big(\bigcap_{j\in\mathbb{Z}_{>0}} S_j\Big) = \lim_{j\to\infty} \mu(S_j)$.

If μ is $\overline{\mathbb{R}}$-valued then μ is:

(iii) ***finite*** if $X \in \mathscr{S}$ and if μ takes values in $\mathbb{R}$;

(iv) ***σ-finite*** if there exists subsets $(S_j)_{j\in\mathbb{Z}_{>0}}$ from $\mathscr{S}$ such that $|\mu(S_j)| < \infty$ for $j \in \mathbb{Z}_{>0}$ and such that $X = \cup_{j\in\mathbb{Z}_{>0}} S_j$.

(v) ***finitely-additive*** if $\mu\Big(\bigcup_{j=1}^{k} S_j\Big) = \sum_{j=1}^{k} \mu(S_j)$ for every finite family $(S_1, \dots, S_k)$ of pairwise disjoint sets from $\mathscr{S}$ whose union is also in $\mathscr{S}$;

(vi) ***countably-additive*** if $\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} S_j\Big) = \sum_{j=1}^{\infty} \mu(S_j)$ for every countable family $(S_j)_{j\in\mathbb{Z}_{>0}}$ of pairwise disjoint sets from $\mathscr{S}$ whose union is also in $\mathscr{S}$.

(vii) ***consistent*** if at most one of ∞ and $-\infty$ is in image(μ). •

Initially, we shall only use the preceding definitions in the case where μ takes values in $\overline{\mathbb{R}}_{\geq 0}$. However, in Sections 5.3.7 and 5.3.8 we shall need to consider the case where μ takes values in $\overline{\mathbb{R}}$.

The following result records some obvious relationships between the preceding concepts.

5.3.2 Proposition (Relationships between properties of functions on subsets) *If* X *is a set, if* $\mathscr{S} \subseteq \mathbf{2}^X$, *and if* $\mu\colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$, *then the following statements hold:*

(i) if $\mu(\emptyset) = 0$ *and if* μ *is countably-subadditive then it is finitely-subadditive;*

(ii) if $\mu(\emptyset) = 0$ *and if* μ *is finitely-additive then it is finitely-subadditive;*

(iii) if $\mu(\emptyset) = 0$ *and if* μ *is countably-additive then it is countably-subadditive;*

(iv) if μ *is countably-additive then it is monotonically increasing;*

(v) if μ *is countably-additive then it is monotonically decreasing;*

(vi) if μ *is finitely-additive then it is monotonic and, moreover,* $\mu(T \setminus S) = \mu(T) - \mu(S)$ *if* $\mu(S) < \infty$.

If μ *takes values in* $\overline{\mathbb{R}}$ *then the following statement holds:*

(vii) if $\mu(\emptyset) = 0$ *and if* μ *is countably-additive then it is finitely-additive.*

If μ *takes values in* $\overline{\mathbb{R}}$ *and if* $\mathscr{S}$ *has the property that* $S \in \mathscr{S}$ *implies that* $X \setminus S \in \mathscr{S}$, *then the following statement holds:*

(viii) if μ *is finitely additive then it is consistent.*

Proof (i) Let $(S_1, \dots, S_k)$ be a finite family of subsets from $\mathscr{S}$. Define $(T_j)_{j\in\mathbb{Z}_{>0}}$ by

$$T_j = \begin{cases} S_j, & j \in \{1, \dots, k\}, \\ \emptyset, & j > k. \end{cases}$$

Then

$$\mu\Big(\bigcup_{j=1}^{k} S_j\Big) = \mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} T_j\Big) \leq \sum_{j=1}^{\infty} \mu(T_j) = \sum_{j=1}^{k} \mu(S_j),$$

since $\mu(\emptyset) = 0$.

The following lemma will be useful in the next two parts of the proof, as well as in various other arguments in this chapter.

1 **Lemma** *Let* X *be a set, let either* $\mathrm{J} = \{1, \dots, \mathrm{m}\}$ *for some* $\mathrm{m} \in \mathbb{Z}_{>0}$ *or* $\mathrm{J} = \mathbb{Z}_{>0}$*, and let* $(\mathrm{S_j})_{\mathrm{j}\in\mathrm{J}}$ *be a finite or countable family of subsets of* X*. Then there exists a family* $(\mathrm{T_j})_{\mathrm{j}\in\mathrm{J}}$ *of subsets of* X *such that*

(i) $\mathrm{T_{j_1}} \cap \mathrm{T_{j_2}} = \emptyset$ *for* $\mathrm{j_1} \neq \mathrm{j_2}$*;*

(ii) $\mathrm{T_j} \subseteq \mathrm{S_j}$, $\mathrm{j} \in \mathrm{J}$*;*

(iii) $\cup_{\mathrm{j}\in\mathrm{J}}\mathrm{T_j} = \cup_{\mathrm{j}\in\mathrm{J}}\mathrm{S_j}$*.*

Moreover, if $\mathrm{S_j} \in \mathscr{A}$, $\mathrm{j} \in \mathbb{Z}_{>0}$*, for an algebra* $\mathscr{A}$ *on* X*, then the sets* $(\mathrm{T_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *can also be chosen to be in* $\mathscr{A}$*.*

Proof Let $j_0 \in J$ and define

$$T'_{j_0} = \bigcup_{j<j_0}(S_{j_0} \cap S_j), \quad T_{j_0} = S_{j_0} \setminus T'_{j_0}.$$

Thus T'_{j_0} is the set of points in S_{j_0} that are already contained in at least one of the "previous" subsets $\{S_j\}_{j<j_0}$, and T_{j_0} is the set of points in S_{j_0} not in one of the sets $\{S_j\}_{j<j_0}$. Thus we immediately have $T_{j_0} \subseteq S_{j_0}$ for each $j_0 \in J$. Let $j_1, j_2 \in J$ be distinct and suppose, without loss of generality, that $j_1 < j_2$. Then, by construction, T_{j_2} contains no points from S_{j_1} and since $T_{j_1} \subseteq S_{j_1}$ our claim follows. Finally, we show that $\cup_{j\in J} T_j = \cup_{j\in J} S_j$. This is clear since T_{j_0} is defined to contain those points from S_{j_0} not already in $S_1, \dots, S_{j_0-1}$.

The last assertion of the lemma follows since the sets T_j, $j \in \mathbb{Z}_{>0}$, are of the form $(X \setminus A) \cap B$ where $A \in \mathscr{A}$ and where B is a union of sets of the form $B_1 \cap B_2$ for $B_1, B_2 \in \mathscr{A}$. Thus $B \in \mathscr{A}$ by Exercise 5.2.1 and so $(X \setminus A) \cap B \in \mathscr{A}$, also by Exercise 5.2.1. ▼

(ii) By the lemma above let $(T_1, \dots, T_m)$ be pairwise disjoint, such that $T_j \subseteq S_j$ for $j \in \{1, \dots, m\}$, and such that $\cup_{j=1}^m T_j = \cup_{j=1}^m S_j$. Then, by finite-additivity,

$$\mu\Big(\bigcup_{k=1}^{m} S_k\Big) = \sum_{k=1}^{m} \mu(T_k).$$

But, for each $k \in \{1, \dots, m\}$, $S_k = S'_k \cup T_k$ and the union is disjoint. Monotonicity of μ gives $\mu(S_j) \geq \mu(T_j)$ for $j \in \{1, \dots, m\}$ which then gives

$$\mu\Big(\bigcup_{k=1}^{m} S_k\Big) = \sum_{k=1}^{m} \mu(T_k) \leq \sum_{k=1}^{m} \mu(S_k),$$

as desired.

(iii) This follows from Lemma 1 just as does part (ii), with only trivial modifications to replace finite-additivity with countable-additivity.

(iv) Let $(S_j)_{j\in\mathbb{Z}_{>0}}$ be a countable family of subsets from $\mathscr{S}$ for which $S_j \subseteq S_{j+1}$, $j \in \mathbb{Z}_{>0}$. For $j \in \mathbb{Z}_{>0}$ define

$$T_j = \begin{cases} S_1, & j = 1, \\ S_j \setminus S_{j-1}, & j > 1. \end{cases}$$

Note that the sets $\{T_j\}_{j\in\mathbb{Z}_{>0}}$ are pairwise disjoint by construction and that

$$\bigcup_{j\in\mathbb{Z}_{>0}} S_j = \bigcup_{j\in\mathbb{Z}_{>0}} T_j.$$

Therefore, by countable-additivity,

$$\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} S_j\Big) = \sum_{j=1}^{\infty} \mu(T_j).$$

But, since $\cup_{j=1}^k T_j = S_k$,

$$\sum_{j=1}^{\infty} \mu(T_j) = \lim_{k\to\infty} \sum_{j=1}^{k} \mu(T_j) = \lim_{k\to\infty} \mu\Big(\bigcup_{j=1}^{k} T_j\Big) = \lim_{k\to\infty} \mu(S_k),$$

which gives

$$\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} S_j\Big) = \lim_{k\to\infty} \mu(S_k),$$

as desired.

(v) Let $(S_j)_{j\in\mathbb{Z}_{>0}}$ be a countable family of sets from $\mathscr{S}$ such that $S_j \supseteq S_{j+1}$, $j \in \mathbb{Z}_{>0}$, and such that $\mu(S_k)$ for some $k \in \mathbb{Z}_{>0}$. Define $(T_j)_{j\in\mathbb{Z}_{>0}}$ by $T_j = S_{j+k}$ so that

$$\bigcap_{j\in\mathbb{Z}_{>0}} S_j = \bigcap_{j\in\mathbb{Z}_{>0}} T_j.$$

Now define $(U_j)_{j\in\mathbb{Z}_{>0}}$ by $U_j = T_1 \setminus T_j$ so that $U_j \subseteq U_{j+1}$ for each $j \in \mathbb{Z}_{>0}$. We also have

$$\bigcup_{j\in\mathbb{Z}_{>0}} U_j = T_1 \setminus \Big(\bigcap_{j\in\mathbb{Z}_{>0}} T_j\Big).$$

By parts (vi) (since $\mu(T_1) < \infty$) and (iv) we then have

$$\begin{aligned}
\mu(T_1) - \mu\Big(\bigcap_{j\in\mathbb{Z}_{>0}} S_j\Big) &= \mu\Big(T_1 \setminus \Big(\bigcap_{j\in\mathbb{Z}_{>0}} S_j\Big)\Big) = \mu\Big(T_1 \setminus \Big(\bigcap_{j\in\mathbb{Z}_{>0}} T_j\Big)\Big) \\
&= \mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} U_j\Big) = \mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} U_j\Big) = \lim_{j\to\infty} \mu(U_j) \\
&= \lim_{j\to\infty} \mu(T_1 \setminus T_j) = \mu(T_1) - \lim_{j\to\infty} \mu(T_j) \\
&= \mu(T_1) - \lim_{j\to\infty} \mu(S_j),
\end{aligned}$$

from which we deduce

$$\mu\Big(\bigcap_{j\in\mathbb{Z}_{>0}} S_j\Big) = \lim_{j\to\infty} \mu(S_j)$$

since $\mu(T_1) < \infty$.

(vi) Let $S, T \in \mathscr{S}$ be such that $S \subseteq T$. Then, by finite-additivity,

$$\mu(S) \leq \mu(S) + \mu(T - S) = \mu(T),$$

as desired. The formula $\mu(T \setminus S) = \mu(T) - \mu(S)$ if $\mu(S) = \infty$ follows immediately from finite-additivity.

(vii) Let $(S_1, \dots, S_k)$ be a finite family of subsets from $\mathscr{S}$. Define $(T_j)_{j \in \mathbb{Z}_{>0}}$ by

$$T_j = \begin{cases} S_j, & j \in \{1, \dots, k\}, \\ \emptyset, & j > k, \end{cases}$$

noting that the family $(T_j)_{j \in \mathbb{Z}_{>0}}$ is pairwise disjoint. Then

$$\mu\Big(\bigcup_{j=1}^{k} S_j\Big) = \mu\Big(\bigcup_{j \in \mathbb{Z}_{>0}} T_j\Big) = \sum_{j=1}^{\infty} \mu(T_j) = \sum_{j=1}^{k} \mu(S_j),$$

since $\mu(\emptyset) = 0$.

(viii) Suppose that there exists sets $S_+, S_- \in \mathscr{S}$ such that $\mu(S_+) = \infty$ and $\mu(S_-) = -\infty$. Then, finite-additivity and the assumption that sets from $\mathscr{S}$ have complements in $\mathscr{S}$ implies that

$$\mu(X) = \mu(S_+) + \mu(X \setminus S_+) = \mu(A_-) + \mu(X \setminus S_-).$$

Since $\mu(S_+) = \infty$ and since $\mu(S_-) = -\infty$ and since the addition $\infty + (-\infty)$ is not defined, we must have

$$\mu(X \setminus S_+) \in \overline{\mathbb{R}} \setminus \{-\infty\}, \quad \mu(X \setminus S_-) \in \overline{\mathbb{R}} \setminus \{\infty\}.$$

Therefore,

$$\mu(X) = \infty, \quad \mu(X) = -\infty,$$

giving a contradiction. ■

The following relationships between finite-additivity, countable-additivity, and monotonicity are also useful.

5.3.3 Proposition (Additivity and monotonicity) *Let* X *be a set with* $\mathscr{A} \subseteq \mathbf{2}^{\mathsf{X}}$ *an algebra on* $\mathscr{A}$*, and let* $\mu_0 \colon \mathscr{A} \to \overline{\mathbb{R}}$ *be consistent, finitely-additive, and have the property that* $\mu_0(\emptyset) = 0$*. The following three statements are equivalent:*

(i) μ_0 *is countably-additive;*

(ii) for every sequence $(\mathsf{A}_\mathsf{j})_{\mathsf{j} \in \mathbb{Z}_{>0}}$ *of subsets from* $\mathscr{A}$ *for which* $\mathsf{A}_\mathsf{j} \subseteq \mathsf{A}_{\mathsf{j}+1}$, $\mathsf{j} \in \mathbb{Z}_{>0}$*, and for which* $\cup_{\mathsf{j} \in \mathbb{Z}_{>0}} \mathsf{A}_\mathsf{j} \in \mathscr{A}$*, it holds that*

$$\mu_0\Big(\bigcup_{\mathsf{j} \in \mathbb{Z}_{>0}} \mathsf{A}_\mathsf{j}\Big) = \lim_{\mathsf{j} \to \infty} \mu_0(\mathsf{A}_\mathsf{j});$$

(iii) for every sequence $(\mathsf{A}_\mathsf{j})_{\mathsf{j} \in \mathbb{Z}_{>0}}$ *of subsets from* $\mathscr{A}$ *for which* $\mathsf{A}_\mathsf{j} \supseteq \mathsf{A}_{\mathsf{j}+1}$, $\mathsf{j} \in \mathbb{Z}_{>0}$*, for which* $\cap_{\mathsf{j} \in \mathbb{Z}_{>0}} \mathsf{A}_\mathsf{j} = \emptyset$*, it holds that* $\lim_{\mathsf{j} \to \infty} \mu_0(\mathsf{A}_\mathsf{j}) = 0$.

Proof (i) $\Longrightarrow$ (ii) Let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of subsets from $\mathscr{A}$ for which $A_j \subseteq A_{j+1}$, $j \in \mathbb{Z}_{>0}$, and for which $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{A}$. Let us denote $A = \cup_{j\in\mathbb{Z}_{>0}} A_j$. Define $B_1 = A_1$ and for $j \geq 2$ define $B_j = A_j \setminus A_{j-1}$. Then the family $(B_j)_{j\in\mathbb{Z}_{>0}}$ is pairwise disjoint and satisfies $\cup_{j\in\mathbb{Z}_{>0}} B_j = A$. The assumed consistency and countable-additivity of μ_0 then gives

$$\mu_0(A) = \sum_{j=1}^{\infty} \mu_0(B_j).$$

Moreover, since $A_k = \cup_{j=1}^{k} B_j$ we have

$$\mu_0(A_k) = \sum_{j=1}^{k} \mu_0(B_j) \quad\Longrightarrow\quad \lim_{k\to\infty} \mu_0(A_k) = \sum_{j=1}^{\infty} \mu_0(B_j) = \mu_0(A),$$

as desired.

(ii) $\Longrightarrow$ (iii) Let us define $B_j = A_{k+j-1}$ for $j \in \mathbb{Z}_{>0}$. Then $\mu_0(B_1) < \infty$ and $\cap_{j\in\mathbb{Z}_{>0}} B_j = \emptyset$. Also define $C_j = B_1 \setminus B_{j+1}$ for $j \in \mathbb{Z}_{>0}$. Then the family of subsets $(C_j)_{j\in\mathbb{Z}_{>0}}$ is in $\mathscr{A}$ and satisfies $C_j \subseteq C_{j+1}$ for each $j \in \mathbb{Z}_{>0}$. Moreover,

$$\bigcup_{j\in\mathbb{Z}_{>0}} C_j = \bigcup_{j\in\mathbb{Z}_{>0}} B_1 \setminus B_{j+1} = B_1 \setminus \bigcap_{j\in\mathbb{Z}_{>0}} B_{j+1} = B_1,$$

using De Morgan's Laws. By assumption we then have

$$\lim_{j\to\infty} \mu_0(C_j) = \mu_0(B_1).$$

Therefore,

$$\lim_{j\to\infty} \mu_0(C_j) = \lim_{j\to\infty} \mu_0(B_1 \setminus B_{j+1}) = \lim_{j\to\infty} (\mu_0(B_1) - \mu_0(B_{j+1})) = \mu_0(B_1),$$

allowing us to conclude that

$$\lim_{j\to\infty} \mu_0(A_j) = \lim_{j\to\infty} \mu_0(B_j) = 0,$$

as desired.

(iii) $\Longrightarrow$ (i) Let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a family of pairwise disjoint sets and denote $A = \cup_{j\in\mathbb{Z}_{>0}} A_j$, supposing that $A \in \mathscr{A}$. For $k \in \mathbb{Z}_{>0}$ define $B_k = A \setminus \cup_{j=1}^{k} A_j$. Then $B_k \supseteq B_{k+1}$ and $\cap_{k\in\mathbb{Z}_{>0}} B_k = \emptyset$. By assumption we then have $\lim_{k\to\infty} \mu_0(B_k) = 0$. We have $A = B_k \cup (\cup_{j=1}^{k} A_j)$ with the union being disjoint. Finite-additivity of μ_0 gives

$$\mu_0(A) = \mu_0(B_k) + \sum_{j=1}^{k} \mu_0(A_j),$$

which gives

$$\mu_0(A) = \lim_{k\to\infty} \mu_0(B_k) + \sum_{j=1}^{\infty} \mu_0(A_j) = \sum_{j=1}^{\infty} \mu_0(A_j),$$

as desired. ■

5.3.2 Outer measures, measures, and their relationship

With the general properties of functions on subsets from the preceding section, we now introduce our first notion of "size" of a subset.

5.3.4 Definition (Outer measure) Let X be a set. An ***outer measure*** on X is a map $\mu^*\colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ with the following properties:

(i) $\mu^*(\emptyset) = 0$;

(ii) μ^* is monotonic;

(iii) μ^* is countably-subadditive. •

5.3.5 Remark (Why are the axioms for an outer measure as they are?) The notion of outer measure is intuitive, in the sense that its properties are included in those that we anticipate a reasonable notion of "size" to possess. What is not immediately clear is that these are the *only* properties that one might demand of our notion of size. This latter matter is difficult to address *a priori*, and indeed is only really addressed by knowing that these are indeed the properties that one uses in the development of the general theory. •

Let us consider some simple examples of outer measures. We shall postpone to Sections 5.4 and 5.5 the presentation of more interesting examples.

5.3.6 Examples (Outer measures)

1. For a set X, the map $\mu^*\colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ defined by $\mu^*(A) = 0$ is an outer measure. We call this the ***zero outer measure***.
2. Let us consider a set X with $\mu^*\colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$\mu^*(A) = \begin{cases} 0, & A = \emptyset, \\ \infty, & A \neq \emptyset. \end{cases}$$

   It is then easy to see that μ^* is an outer measure.
3. For a set X define $\mu^*\colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\mu^*(A) = \begin{cases} \operatorname{card}(A), & \operatorname{card}(A) < \infty, \\ \infty, & \text{otherwise}. \end{cases}$$

   It is easy to verify that μ^* is an outer measure.
4. For a set X define $\delta_x^*\colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\delta_x^*(A) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

   One can easily see that δ_x^* is indeed an outer measure. •

The notion of outer measure is a nice one in that it allows the measurement of size for any subset of the set X. However, it turns out that some outer measures lack an important property. Namely, there are outer measures μ^* (namely, the Lebesgue outer measure of Definition 5.4.1) that lack the property that, if $S, T \subseteq X$ are *disjoint*, then $\mu^*(S \cup T) = \mu^*(S) + \mu^*(T)$. Upon reflection, we hope the reader can see that this is indeed a property one would like any notion of size to possess. In order to ensure that this property is satisfied, it turns out that one needs to restrict oneself to measuring a subset of the collection of all sets. It is here where the notions of algebras and σ-algebras come into play.

5.3.7 Definition (Measure, measure space) Let X be a set and let $\mathscr{S} \subseteq \mathbf{2}^X$. A ***finitely-additive measure*** on $\mathscr{S}$ is a map $\mu\colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$ with the following properties:

(i) $\mu(\emptyset) = 0$;

(ii) μ is finitely-additive.

A ***countably-additive measure***, or simply a ***measure***, on $\mathscr{S}$ is a map $\mu\colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$ with the following properties:

(iii) $\mu(\emptyset) = 0$;

(iv) μ is countably-additive.

A triple $(X, \mathscr{A}, \mu)$ is called a ***measure space*** is $\mathscr{A}$ is a σ-algebra on X and if μ is a countably-additive measure on $\mathscr{A}$. •

Just as we are primarily interested in σ-algebras in preference to algebras, we are also primarily interested in countably-additive measures in preference to finitely-additive measures. However, finitely-additive measures will come up, usually in the course of a construction of a countably-additive measure.

5.3.8 Remark (Why are the axioms for a measure as they are?) Again, it is not perfectly evident why a measure has the stated properties. In particular, the conditions that (1) a measure space involves a σ-algebra and that (2) a measure be countably-additive seem like they ought to admit many viable alternatives. Why not allow a measure space to be defined using *any* collection of subsets? Why not finite-additivity? finite-subadditivity? countable-subadditivity? The reasons to restrict to a σ-algebra (possibly) smaller than the collection of all subsets will be made clear shortly. As concerns *countable*-additivity, the reasons for this are much like they are for the countability conditions for σ-algebras; countability is what we want here since we are after nice behaviour of our constructions with sequential operations. The requirement of disjointness in the definition is not so easy to understand. Indeed, in our definition of an outer measure in Definition 5.3.4 we relaxed this, and possibly the definition of an outer measure seems like the one that we should really be interested in. However, it is not, although the reasons for this will only be made clear as we go along. •

Let us give some simple examples of measures.

5.3.9 Examples (Measures)

1. For a measurable space $(X, \mathscr{A})$, the map $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ defined by $\mu^*(A) = 0$ is a measure. We call this the ***outer measure***.
2. For a measurable space $(X, \mathscr{A})$ define $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\mu(A) = \begin{cases} 0, & A = \emptyset, \\ \infty, & A \neq \emptyset. \end{cases}$$

This defines a measure on $(X, \mathscr{A})$.

3. If $(X, \mathscr{A})$ is a measurable space then define $\mu_\Sigma\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\mu_\Sigma(A) = \begin{cases} \operatorname{card}(A), & \operatorname{card}(A) < \infty, \\ \infty, & \text{otherwise.} \end{cases}$$

One may verify that this defines a measure for the measurable space $(X, \mathscr{A})$ called the ***counting measure***.

4. If $(X, \mathscr{A})$ is a measurable space and if $x \in X$ we define $\delta_x\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\delta_x(A) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

One may verify that this defines a measure and is called the ***point mass*** concentrated at x.

5. On the algebra $\mathscr{J}(\mathbb{R}^n)$ of Jordan measurable subsets of $\mathbb{R}^n$ the map $\rho\colon \mathscr{J}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ of Definition 5.1.6 is a finitely-additive measure. This follows from Proposition 5.1.8. •

Let us give some properties of measures that follow more or less directly from the definitions.

5.3.10 Proposition (Properties of measures) *For a set* X*, a collection of subsets* $\mathscr{S} \subseteq 2^X$*, and a measure* μ *on* $\mathscr{S}$*, the following statements hold:*

(i) *μ is finitely-additive;*

(ii) *μ is monotonic and $\mu(T \setminus S) = \mu(T) - \mu(S)$ if $\mu(S) < \infty$;*

(iii) *μ is countably-subadditive;*

(iv) *μ is monotonically increasing;*

(v) *μ is monotonically decreasing.*

Proof (i) This follows immediately from Proposition 5.3.2(vii).

(ii) This follows from Proposition 5.3.2(vi) and part (i).

(iii) This follows from Proposition 5.3.2(iii).

(iv) This follows from Proposition 5.3.2(iv).

(v) This follows from Proposition 5.3.2(v). ■

Now let us examine the relationships between outer measure and measure. Let us begin with something elementary, given what we already know.

5.3.11 Proposition (When are measures outer measures?) *If $(X, \mathscr{A}, \mu)$ is a measure space then μ is an outer measure if and only if $\mathscr{A} = 2^X$.*

Proof This follows immediately from Proposition 5.3.10. ■

Since the outer measures in the examples are all actually measures, this leads one to the following line of questioning.

1. Are all outer measures measures?

2. Given a measure space $(X, \mathscr{A}, \mu)$ does there exist an outer measure μ^* on X for which $\mu = \mu^*|\mathscr{A}$?

We shall see in Corollary 5.3.29 that the answer to the second question is, "Yes." The answer to the first question is, "No," but we will have to wait until Section 5.4 (in particular, Example 5.4.3) to see an example of an outer measure that is not a measure. The key issue concerning whether an outer measure is a measure hinges on the following characterisation of a distinguished class of subsets of a set with an outer measure.

5.3.12 Definition (Measurable subsets for an outer measure) If μ^* is an outer measure on a set X, a subset $A \subseteq X$ is ***μ^*-measurable*** if

$$\mu^*(S) = \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A))$$

for all $S \subseteq X$. The set of μ^*-measurable subsets is denoted by $\mathscr{M}(X, \mu^*)$. •

Note that an outer measure is finitely-subadditive by Proposition 5.3.2(i). Thus we always have

$$\mu^*(S) \leq \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)).$$

Therefore, a set A is *not* μ^*-measurable then we have

$$\mu^*(S) > \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)).$$

The definition of μ^*-measurability looks like it provides a "reasonable" property of a subset A: that the outer measure of a set S should be the outer measure of the points in S that are in A plus the outer measure of the points in S that are not in A. In Figure 5.2 we attempt to depict what is going on. What is not so obvious is that not all subsets need be μ^*-measurable. In the examples of outer measures in Example 5.3.6 above, they all turn out to be measures. It is only when we get

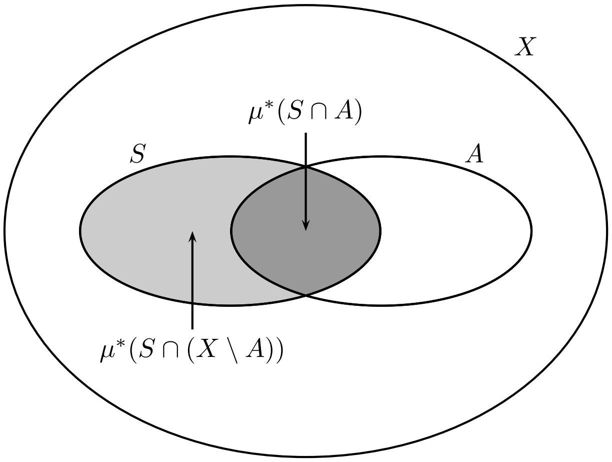


Figure 5.2 The notion of a μ^*-measurable set

to the more sophisticated construction of the Lebesgue measure in Section 5.4 that we see that nonmeasurable sets exist. Indeed, it is precisely in the constructions of Section 5.4 that the general ideas we are presently discussing were developed.

For the purposes of our present development, the following theorem is important in that it gives a natural passage from an outer measure to a measure space.

5.3.13 Theorem (Outer measures give measure spaces) *If μ^* is an outer measure on a set* X *then* $(\mathsf{X}, \mathscr{M}(\mathsf{X},\mu^*), \mu^*|\mathscr{M}(\mathsf{X},\mu^*))$ *is a measure space.*

Proof Let us first show that $X \in \mathscr{M}(X,\mu^*)$. Let $S \in \mathbf{2}^X$ and note that

$$\mu^*(S \cap X) + \mu^*(S \cap (X \setminus X)) = \mu^*(S)$$

since $\mu^*(\emptyset) = 0$.

Now let us show that if $A \in \mathscr{M}(X,\mu^*)$ then $X \setminus A \in \mathscr{M}(X,\mu^*)$. This follows since

$$\mu^*(S \cap (X \setminus A)) + \mu^*(S \cap (X \setminus (X \setminus A))) = \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)) = \mu^*(S).$$

Next we show that if $A_1, \dots, A_n \in \mathscr{M}(X,\mu^*)$ then $\cup_{j=1}^n A_j \in \mathscr{M}(X,\mu^*)$. This will follow by a trivial induction if we can prove it for $n = 2$. Thus we let $A_1, A_2 \in \mathscr{M}(X,\mu^*)$, $S \subseteq X$, and compute

$$\begin{aligned}
&\mu^*(S \cap (A_1 \cup A_2)) + \mu^*(S \cap (X \setminus (A_1 \cup A_2)))\\
&\quad = \mu^*((S \cap (A_1 \cup A_2)) \cap A_1) + \mu^*((S \cap (A_1 \cup A_2)) \cap (X \setminus A_1)) + \mu^*(S \cap (X \setminus (A_1 \cup A_2)))\\
&\quad = \mu^*(S \cap A_1) + \mu^*(S \cap (X \setminus A_1) \cap A_2) + \mu^*(S \cap (X \setminus A_1) \cap (X \setminus A_2))\\
&\quad = \mu^*(S \cap A_1) + \mu^*(S \cap (X \setminus A_1)) = \mu^*(S).
\end{aligned}$$

In going from the first line to the second line we have used the fact that $A_1 \in \mathscr{M}(X,\mu^*)$. In going from the second line to the third line we have used some set theoretic identities for union and intersection that can be easily verified, e.g., by using Propositions 1.1.4 and 1.1.5. In going from the third line to the fourth line we have used the fact that $A_2 \in \mathscr{M}(X,\mu^*)$.

Next we show that property (vi) of Definition 5.2.1 holds. Thus we let $(A_j)_{j\in\mathbb{Z}_{>0}} \subseteq \mathscr{M}(X,\mu^*)$. To show that $\cup_{j\in\mathbb{Z}_{>0}} A_j \in \mathscr{M}(X,\mu^*)$ we may without loss of generality suppose that the sets $(A_j)_{j\in\mathbb{Z}_{>0}}$ are disjoint. Indeed, if they are not then we may replace their union with the union of the sets

$$\begin{aligned}
\tilde{A}_1 &= A_1\\
\tilde{A}_2 &= A_2 \cap (X \setminus A_1)\\
&\ \ \vdots\\
\tilde{A}_j &= A_j \cap (X \setminus A_1) \cap \dots \cap (X \setminus A_{j-1})\\
&\ \ \vdots
\end{aligned}$$

where the collection $(\tilde{A}_j)_{j\in\mathbb{Z}_{>0}}$ is disjoint. First we claim that under this assumption that $(A_j)_{j\in\mathbb{Z}_{>0}} \subseteq \mathscr{M}(X,\mu^*)$ is disjoint we have

$$\mu^*(S) = \sum_{j=1}^n \mu^*(S \cap A_j) + \mu^*\Big(S \cap \Big(\bigcap_{j=1}^n (X \setminus A_j)\Big)\Big). \tag{5.1}$$

We prove this by induction. For $n = 1$ the claim follows since $A_1 \in \mathscr{M}(X, \mu^*)$. Now suppose the claim true for $n = k$ and compute

$$\begin{aligned}
&\mu^*\Big(S \cap \Big(\bigcap_{j=1}^{k} (X \setminus A_j)\Big)\Big) \\
&= \mu^*\Big(S \cap \Big(\bigcap_{j=1}^{k} (X \setminus A_j)\Big) \cap A_{k+1}\Big) + \mu^*\Big(S \cap \Big(\bigcap_{j=1}^{k} (X \setminus A_j)\Big) \cap (X \setminus A_{k+1})\Big) \\
&= \mu^*(S \cap A_{k+1}) + \mu^*\Big(S \cap \Big(\bigcap_{j=1}^{k+1} (X \setminus A_j)\Big)\Big),
\end{aligned}$$

so establishing (5.1) after an application of the induction hypothesis. In the first line we use the fact that $A_{k+1} \in \mathscr{M}(X, \mu^*)$ and in the second line we have used the fact that the set $(A_j)_{j \in \mathbb{Z}_{>0}}$ are disjoint.

By monotonicity of outer measures we have

$$\begin{aligned}
&\mu^*(S) \geq \sum_{j=1}^{n} \mu^*(S \cap A_j) + \mu^*\Big(S \cap \Big(\bigcap_{j=1}^{\infty} (X \setminus A_j)\Big)\Big) \\
\implies\quad &\mu^*(S) \geq \sum_{j=1}^{n} \mu^*(S \cap A_j) + \mu^*\Big(S \cap \Big(X \setminus \bigcup_{j=1}^{\infty} A_j\Big)\Big) \\
\implies\quad &\mu^*(S) \geq \sum_{j=1}^{\infty} \mu^*(S \cap A_j) + \mu^*\Big(S \cap \Big(X \setminus \bigcup_{j=1}^{\infty} A_j\Big)\Big) \qquad (5.2) \\
\implies\quad &\mu^*(S) \geq \mu^*\Big(S \cap \Big(\bigcup_{j=1}^{\infty} A_j\Big)\Big) + \mu^*\Big(S \cap \Big(X \setminus \bigcup_{j=1}^{\infty} A_j\Big)\Big).
\end{aligned}$$

In the first line we have used (5.1) along with monotonicity of outer measures. In the second line we have used a simple set theoretic identity. In the third line we have simply taken the limit of a bounded monotonically increasing sequence of numbers. In the fourth line we have used countable-subadditivity of outer measures. This then gives

$$\mu^*(S) \geq \mu^*\Big(S \cap \Big(\bigcup_{j=1}^{\infty} A_j\Big)\Big) + \mu^*\Big(S \cap \Big(X \setminus \bigcup_{j=1}^{\infty} A_j\Big)\Big) \geq \mu^*(S),$$

by another application countable-subadditivity of outer measures. It therefore follows that $\cup_{j \in \mathbb{Z}_{>0}} A_j \in \mathscr{M}(X, \mu^*)$, as was to be shown.

The next thing we show is that $\mu \triangleq \mu^* | \mathscr{M}(X, \mu^*)$ is a measure on $(X, \mathscr{M}(X, \mu^*))$. Since

$$\mu^*(S) = \mu^*(S \cap \emptyset) + \mu^*(S \cap X) = \mu^*(\emptyset) + \mu^*(S),$$

for every $S \in \mathbf{2}^X$ it follows that $\mu(\emptyset) = \mu^*(\emptyset) = 0$. Now let $(A_j)_{j \in \mathbb{Z}_{>0}}$ be a collection of disjoint sets in $\mathscr{M}(X, \mu^*)$. We have

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \mu^*\Big(\bigcup_{j=1}^{\infty} A_j\Big) \geq \sum_{j=1}^{\infty} \mu^*(A_j) + 0,$$

using (5.2) with $S = \cup_{j=1}^{\infty} A_j$. By monotonicity of outer measures we also have

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \le \sum_{j=1}^{\infty} \mu^*(A_j),$$

and so μ is countably-additive. ■

The theorem immediately has the following corollary which helps to clarify the relationship between measures and outer measures.

5.3.14 Corollary (An outer measure is a measure if and only if all subsets are measurable) *If μ^* is an outer measure on* X *then* $(\mathsf{X}, \mathbf{2}^{\mathsf{X}}, \mu^*)$ *is a measure space if and only if every subset of* X *is μ^*-measurable.*

Proof From Theorem 5.3.13 it follows that $(X, \mathbf{2}^X, \mu^*)$ is a measure space if $\mathscr{M}(X, \mu^*) = \mathbf{2}^X$. For the converse, suppose that $A \subseteq X$ is not μ^*-measurable. Then there exists a set $S \subseteq X$ such that

$$\mu^*(S) \neq \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)).$$

However, since $S = (S \cap A) \cup (S \cap (X \setminus A))$ this prohibits μ^* from being a measure since, if it were a measure, we would have

$$\mu^*(S) = \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)).$$ ■

Thus the existence of nonmeasurable sets is exactly the obstruction to an outer measure being a measure. Said otherwise, if we wish for an outer measure to behave like a measure— i.e., have the property that

$$\mu^*\Big(\bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) = \sum_{j=1}^{\infty} \mu^*(A_j)$$

for a family $(A_j)_{j \in \mathbb{Z}_{>0}}$ of disjoint sets—then the sacrifice we have to make is that we possibly restrict the sets which we apply the outer measure to.

The following notions are also sometimes useful.

5.3.15 Definition (Continuous measure, discrete measure) Let $(X, \mathscr{A}, \mu)$ be a measure space for which $\{x\} \in \mathscr{A}$ for every $x \in X$. The measure μ is

(i) ***continuous*** if $\mu(\{x\}) = 0$ for every $x \in X$ and

(ii) ***discrete*** if there exists a countable subset $D \in \mathscr{A}$ such that $\mu(X \setminus D) = 0$. •

Let us consider how these various properties show up in our simple examples of measure spaces.

5.3.16 Examples (Properties of measures)

1. We consider the measure space $(X, \mathscr{A}, \mu)$ where $\mu(\emptyset) = 0$ and $\mu(A) = \infty$ for all nonempty measurable sets. This measure space is σ-finite if and only if $X = \emptyset$, is continuous if and only if $X = \emptyset$, and is discrete if and only if X is countable.

2. Let us consider a measurable space $(X, \mathscr{A})$ and for simplicity assume that $\{x\} \in \mathscr{A}$ for every $x \in X$. The counting measure is σ-finite if and only if X is countable, is not continuous, and is discrete if and only if X is countable.
3. For a measurable space $(X, \mathscr{A})$ the point mass measure δ_x is σ-finite if and only if X is a countable union of measurable sets, is not continuous, and is discrete if and only if there exists a countable set $D \in \mathscr{A}$ such that $x \notin D$. •

Let us close this section by introducing an important piece of lingo.

5.3.17 Notation (Almost everywhere, a.e.) Let $(X, \mathscr{A}, \mu)$ be a measure space. A property P of the set X holds ***μ-almost everywhere*** (***μ-a.e.***) if there exists a set $A \subseteq X$ for which $\mu(A) = 0$, and such that P holds for all $x \in X \setminus A$. If μ is understood, then we may simply write ***almost everywhere*** (***a.e.***). Some authors use "p.p." after the French "presque partout." Lebesgue, after all, was French. •

Let us finally show that the restriction of a measure to a subset makes sense if the subset is measurable.

5.3.18 Proposition (Restriction of measure to measurable subsets) *Let* $(\mathrm{X}, \mathscr{A}, \mu)$ *be a measure space, let* $\mathrm{A} \in \mathscr{A}$*, let* $(\mathrm{A}, \mathscr{A}_{\mathrm{A}})$ *be the measurable space of Proposition 5.2.6, and define* $\mu_{\mathrm{A}} \colon \mathscr{A}_{\mathrm{A}} \to \overline{\mathbb{R}}_{\geq 0}$ *by* $\mu_{\mathrm{A}}(\mathrm{A} \cap \mathrm{B}) = \mu(\mathrm{A} \cap \mathrm{B})$*. Then* $(\mathrm{A}, \mathscr{A}_{\mathrm{A}}, \mu_{\mathrm{A}})$ *is a measure space.*

Proof It is clear that $\mu_A(\emptyset) = 0$. Also let $(B_j \cap A)_{j \in \mathbb{Z}_{>0}}$ be a countable family of disjoint sets in $\mathscr{A}_A$. Since $B_j \cap A \in \mathscr{A}$ for $j \in \mathbb{Z}_{>0}$ this immediately implies that

$$\mu_A\Big(\bigcup_{j \in \mathbb{Z}_{>0}} B_j \cap A\Big) = \sum_{j=1}^{\infty} \mu_A(B_j \cap A),$$

thus showing that μ_A is a measure on $(A, \mathscr{A}_A)$. ■

5.3.3 Complete measures and completions of measures

In this section we consider a rather technical property of measure spaces, but one that does arise on occasion. It is a property that is at the same time (occasionally) essential and (occasionally) bothersome. This is especially true of the Lebesgue measure we consider in Sections 5.4 and 5.5. We shall point out instances of both of these attributes as we go along.

First we give the definition.

5.3.19 Definition (Complete measure) A measure space $(X, \mathscr{A}, \mu)$ is ***complete*** if for every pair of sets A and B with the properties that $A \subseteq B$, $B \in \mathscr{A}$, and $\mu(B) = 0$, we have $A \in \mathscr{A}$. •

Note that completeness has the interpretation that every subset of a set of measure zero should itself be in the set of measurable subsets, and have measure zero. This seems like a reasonable restriction, but it is one that is not met in certain common examples (see ***missing stuff***). In cases where we have a measure space that is not complete one can simply add some sets to the collection of measurable sets that make the resulting measure space complete. This is done as follows.

5.3.20 Definition (Completion of a measure space) For a measure space $(X, \mathscr{A}, \overline{\mu})$ the ***completion*** $\mathscr{A}$ ***under*** $\boldsymbol{\mu}$ is the collection $\mathscr{A}_\mu$ of subsets $A \subseteq X$ for which there exists $L, U \in \mathscr{A}$ such that $L \subseteq A \subseteq U$ and $\mu(U \setminus L) = 0$. Define $\overline{\mu}\colon \mathscr{A}_\mu \to \overline{\mathbb{R}}_{\geq 0}$ by $\overline{\mu}(A) = \mu(U) = \mu(L)$ where U and L are any sets satisfying $L \subseteq A \subseteq U$ and $\mu(U \setminus L) = 0$. The triple $(X, \mathscr{A}_\mu, \overline{\mu})$ is the ***completion*** of $(X, \mathscr{A}, \mu)$. •

The completion of a measure space is a complete measure space, as we now show.

5.3.21 Proposition (The completion of a measure space is complete) *If* $(\mathsf{X}, \mathscr{A}_\mu, \overline{\mu})$ *is the completion of* $(\mathsf{X}, \mathscr{A}, \mu)$ *then* $(\mathsf{X}, \mathscr{A}_\mu, \overline{\mu})$ *is a complete measure space for which* $\mathscr{A} \subseteq \mathscr{A}_\mu$*.*

Proof If $A \in \mathscr{A}$ then $A \subseteq A \subseteq A$ so that $A \in \mathscr{A}_\mu$. In particular, $X \in \mathscr{A}_\mu$. Note that $L \subseteq A \subseteq U$ and $\mu(U \setminus L) = 0$ implies that $(X \setminus U) \subseteq (X \setminus A) \subseteq (X \setminus L)$ and that $\mu((X \setminus L) \setminus (X \setminus U)) = 0$, thus showing that $X \setminus A \in \mathscr{A}_\mu$. Now let $(A_j)_{j\in\mathbb{Z}_{>0}} \subseteq \mathscr{A}_\mu$ and let $(L_j)_{j\in\mathbb{Z}_{>0}}$ and $(U_j)_{j\in\mathbb{Z}_{>0}}$ satisfy

$$L_j \subseteq A_j \subseteq U_j, \qquad \mu(U_j \setminus L_j) = 0, \qquad j \in \mathbb{Z}_{>0}. \tag{5.3}$$

A direct computation shows that

$$\bigcup_{j\in\mathbb{Z}_{>0}} L_j \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} A_j \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} U_j, \qquad \mu\Big(\Big(\bigcup_{j\in\mathbb{Z}_{>0}} U_j\Big) \setminus \Big(\bigcup_{j\in\mathbb{Z}_{>0}} L_j\Big)\Big) \leq \sum_{j=1}^{\infty} \mu(U_j \setminus L_j) = 0.$$

This shows that $\mathscr{A}_\mu$ is a σ-algebra.

Note that $\overline{\mu}(\emptyset) = 0$. Also let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a collection of disjoint subsets in $\mathscr{A}_\mu$ and take $(L_j)_{j\in\mathbb{Z}_{>0}}$ and $(U_j)_{j\in\mathbb{Z}_{>0}}$ to satisfy (5.3). Note that the sets $(L_j)_{j\in\mathbb{Z}_{>0}}$ are disjoint. From the definition of $\overline{\mu}$ it then follows that $\overline{\mu}$ is countably-additive. It remains to show that $(X, \mathscr{A}_\mu, \overline{\mu})$ is complete. If $A \in \mathscr{A}_\mu$ and $B \subseteq X$ satisfy $B \subseteq A$ and $\overline{\mu}(A) = 0$ then, since $A \in \mathscr{A}_\mu$, we have $U \in \mathscr{A}$ so that $A \subseteq U$ and $\mu(U) = 0$. Taking $L = \emptyset$ we have $L \subseteq B \subseteq U$ and $\mu(U \setminus L) = 0$, showing that $B \in \mathscr{A}_\mu$, as desired. ■

It turns out that the construction in Theorem 5.3.13 of a measure space from an outer measure yields a complete measure space.

5.3.22 Proposition (Completeness of measure space constructed from outer measures) *If* μ^* *is an outer measure on a set* X *then* $(\mathsf{X}, \mathscr{M}(\mathsf{X}, \mu^*), \mu^*|\mathscr{M}(\mathsf{X}, \mu^*))$ *is a complete measure space.*

Proof From Theorem 5.3.13 we need only prove completeness. We let $\mu = \mu^*|\mathscr{M}(X, \mu^*)$. Let $B \in \mathscr{M}(X, \mu^*)$ and let $A \subseteq B$. For $S \in \mathbf{2}^X$ we then have

$$\begin{aligned}\mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)) &\leq \mu^*(S \cap B) + \mu^*(S \cap (X \setminus A))\\ &= 0 + \mu^*(S \cap (X \setminus A)) \leq \mu^*(S),\end{aligned}$$

using the fact that $\mu^*(S \cap B) \leq \mu^*(B) = 0$ and monotonicity of outer measures. By countable-subadditivity of μ^* we have

$$\mu^*(S) \leq \mu^*(S \cap A) + \mu^*(S \cap (X \setminus A)),$$

and so it follows that $A \in \mathscr{M}(X, \mu^*)$. ■

Let us finally show that completeness is preserved by restriction.

5.3.23 Proposition (The restriction of a complete measure is complete) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a complete measure space then the measure space* $(\mathsf{A}, \mathscr{A}_{\mathsf{A}}, \mu_{\mathsf{A}})$ *of Proposition 5.3.18 is complete.*

Proof If $B \cap A \in \mathscr{A}_A$ satisfies $\mu_A(B \cap A) = 0$ then $\mu(B \cap A) = 0$. Therefore, by completeness of μ, if $C \subseteq (B \cap A)$ it follows that $\mu_A(C) = 0$. ■

5.3.4 Outer and inner measures associated to a measure

In this section we continue our exploration of the relationship between outer measure and measure, now going from a measure to an outer measure. We begin with a discussion of ways in which one may generate an outer measure from other data.

5.3.24 Proposition (Outer measure generated by a collection of subsets) *Let* X *be a set, let* $\mathscr{S} \subseteq \mathbf{2}^{\mathsf{X}}$ *have the property that* $\emptyset \in \mathscr{S}$*, and let* $\mu_0 \colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$ *have the property that*

$$\inf\{\mu_0(\mathsf{S}) \mid \mathsf{S} \in \mathscr{S}\} = 0.$$

If we define $\mu^* \colon \mathbf{2}^{\mathsf{X}} \to \overline{\mathbb{R}}_{\geq 0}$ *by*

$$\mu^*(\mathsf{A}) = \inf\Big\{\sum_{\mathsf{j}=1}^{\infty} \mu_0(\mathsf{S}_\mathsf{j}) \Bigm| \mathsf{A} \subseteq \bigcup_{\mathsf{j} \in \mathbb{Z}_{>0}} \mathsf{S}_\mathsf{j},\ \mathsf{S}_\mathsf{j} \in \mathscr{S},\ \mathsf{j} \in \mathbb{Z}_{>0}\Big\},$$

then μ^* *is an outer measure on* X*. Moreover, if* $\mathscr{S}$ *is an algebra on* X *and if* μ_0 *is a countably-additive measure, then* $\mu^*(\mathsf{S}) = \mu_0(\mathsf{S})$ *for every* $\mathsf{S} \in \mathscr{S}$*.*

Proof First let us show that $\mu^*(\emptyset) = 0$. Let $\epsilon \in \mathbb{R}_{>0}$. By hypothesis there exists $S \in \mathscr{S}$ such that $\mu_0(S) \leq \epsilon$, and since $\emptyset \subseteq S$ we have

$$\mu^*(\emptyset) \leq \mu_0(S) \leq \epsilon.$$

As this holds for every $\epsilon \in \mathbb{R}_{>0}$ it follows that $\mu^*(\emptyset) = 0$. That $\mu^*(A) \leq \mu^*(B)$ if $A \subseteq B$ is clear. Now let $(A_j)_{j \in \mathbb{Z}_{>0}}$ be a countable family of subsets of X. If $\sum_{j=1}^{\infty} \mu^*(A_j) = \infty$ then the property of countable-subadditivity holds for the family $(A_j)_{j \in \mathbb{Z}_{>0}}$. Thus suppose that $\sum_{j=1}^{\infty} \mu^*(A_j) < \infty$ and let $\epsilon \in \mathbb{R}_{>0}$. For each $j \in \mathbb{Z}_{>0}$ let $(S_{jk})_{k \in \mathbb{Z}_{>0}}$ be a family of subsets from $\mathscr{S}$ with the properties that $A_j \subseteq \cup_{k \in \mathbb{Z}_{>0}} S_{jk}$ and

$$\sum_{k=1}^{\infty} \mu_0(S_{jk}) < \mu^*(A_j) + \frac{\epsilon}{2^j},$$

this being possible by definition of μ^*. Then

$$\bigcup_{j \in \mathbb{Z}_{>0}} A_j \subseteq \bigcup_{j,k \in \mathbb{Z}_{>0}} S_{jk} \quad \Longrightarrow \quad \mu^*\Big(\bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) \leq \mu^*\Big(\bigcup_{j,k \in \mathbb{Z}_{>0}} S_{jk}\Big).$$

Also

$$\bigcup_{j,k \in \mathbb{Z}_{>0}} S_{jk} \subseteq \bigcup_{j,k \in \mathbb{Z}_{>0}} S_{jk} \quad \Longrightarrow \quad \mu^*\Big(\bigcup_{j,k \in \mathbb{Z}_{>0}} S_{jk}\Big) \leq \sum_{j,k=1}^{\infty} \mu_0(S_{jk}) < \sum_{j=1}^{\infty} \mu^*(A_j) + \epsilon,$$

using the fact that $\sum_{j=1}^{\infty} \frac{1}{2^j} = 1$ (see Example 2.4.2–**??**). From this we conclude that

$$\mu^*\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) \le \sum_{j=1}^{\infty} \mu^*(A_j)$$

since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary in the above development. Thus shows that μ^* is indeed an outer measure.

Now we prove the final assertion. Let $S \in \mathscr{S}$. Since $S \subseteq S$ we have $\mu^*(S) \le \mu_0(S)$. Now let $(S_j)_{j\in\mathbb{Z}_{>0}}$ be a family of subsets such that $S \subseteq \cup_{j\in\mathbb{Z}_{>0}} S_j$. Then we define

$$\begin{aligned}
\tilde{S}_1 &= S_1\\
\tilde{S}_2 &= S_2 \cap (X \setminus S_1)\\
&\ \ \vdots\\
\tilde{S}_j &= S_j \cap (X \setminus S_1) \cap \cdots \cap (X \setminus S_{j-1})\\
&\ \ \vdots
\end{aligned}$$

noting that the family of sets $(\tilde{S}_j)_{j\in\mathbb{Z}_{>0}}$ is in $\mathscr{S}$ since $\mathscr{S}$ is an algebra. Moreover, by construction, the sets $(\tilde{S}_j)_{j\in\mathbb{Z}_{>0}}$ are pairwise disjoint and satisfy

$$\bigcup_{j\in\mathbb{Z}_{>0}} S_j = \bigcup_{j\in\mathbb{Z}_{>0}} \tilde{S}_j.$$

Since $\tilde{S}_j \subseteq S_j$ we have

$$\sum_{j=1}^{\infty} \mu_0(\tilde{S}_j) \le \sum_{j=1}^{\infty} \mu_0(S_j).$$

Now, for each $j \in \mathbb{Z}_{>0}$, define $T_j = S \cap \tilde{S}_j$, noting that $T_j \in \mathscr{S}$ since $\mathscr{S}$ is an algebra. Note that $S = \cup_{j\in\mathbb{Z}_{>0}} T_j$. Moreover, by construction the family of sets $(T_j)_{j\in\mathbb{Z}_{>0}}$ is disjoint. Since μ_0 is a measure we have

$$\mu_0(S) = \mu_0\Big(\bigcup_{j\in\mathbb{Z}_{>0}} \tilde{T}_j\Big) = \sum_{j=1}^{\infty} \mu_0(\tilde{T}_j).$$

Since $T_j \subseteq \tilde{S}_j$ we have

$$\sum_{j=1}^{\infty} \mu_0(T_j) \le \sum_{j=1}^{\infty} \mu_0(\tilde{S}_j),$$

giving

$$\sum_{j=1}^{\infty} \mu_0(S_j) \ge \mu_0(S).$$

This allows us to conclude that $\mu^*(S) \ge \mu_0(S)$, and so $\mu^*(S) = \mu_0(S)$, as desired. ■

The outer measure of the preceding proposition has a name.

5.3.25 Definition (Outer measure generated by a collection of sets and a function on those sets) Let X be a set, let $\mathscr{S} \subseteq \mathbf{2}^X$ have the property that $\emptyset \in \mathscr{S}$, and let $\mu_0 \colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$ have the property that

$$\inf\{\mu_0(S) \mid S \in \mathscr{S}\} = 0.$$

The outer measure $\mu^* \colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ defined in Proposition 5.3.24 is the outer measure ***generated*** by the pair $(\mathscr{S}, \mu_0)$. •

Let us give an application of the preceding constructions. A common construction with measures is the extension of a $\overline{\mathbb{R}}_{\geq 0}$-valued function on a collection of subsets to a measure on the σ-algebra generated by the subsets. There are a number of such statements, but the one that we will use is the following.

5.3.26 Theorem (Hahn–Kolmogorov[5] Extension Theorem) *Let* X *be a set, let* $\mathscr{A}$ *be an algebra on* X*, and let* $\mu_0 \colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ *be a σ-finite measure on* $\mathscr{A}$*. Then there exists a unique measure μ on $\sigma(\mathscr{S})$ such that* $\mu(\mathrm{A}) = \mu_0(\mathrm{A})$ *for every* $\mathrm{A} \in \mathscr{A}$*.*

Proof First let us assume that $\mu_0(X) < \infty$. Let $\mu^* \colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ be the outer measure generated by $\mathscr{A}$ and μ_0 as in Proposition 5.3.24. Then Proposition 5.3.24 ensures that $\mu^*(A) = \mu_0(A)$ for every $A \in \mathscr{A}$.

We wish to show that $\mu^*|\sigma(\mathscr{A})$ is a measure. To do this we define $d_{\mu^*} \colon \mathbf{2}^X \times \mathbf{2}^X \to \mathbb{R}_{\geq 0}$ by

$$d_{\mu^*}(S, T) = \mu^*(S \triangle T),$$

recalling from Section 1.1.2 the definition of the symmetric complement $\triangle$. We clearly have $d_{\mu^*}(S, T) = d_{\mu^*}(T, S)$ for every $S, T \subseteq X$. Since μ^* is an outer measure we have

$$\begin{aligned} d_{\mu^*}(S, U) = \mu^*(S \triangle U) &\leq \mu^*((S \triangle T) \cup (T \triangle U)) \\ &\leq \mu^*(S \triangle T) + \mu^*(T \triangle U) = d_{\mu^*}(S, T) + d_{\mu^*}(T, U) \end{aligned}$$

for every $S, T, U \subseteq X$, using Exercise 1.1.2. Thus d_{μ^*} is a semimetric on $\mathbf{2}^X$. Moreover, $d_{\mu^*}(S, T) = 0$ if and only if $\mu^*(S - T) = 0$ and $\mu^*(T - S) = 0$. Thus the implication

$$d_{\mu^*}(S, T) = 0 \quad \Longrightarrow \quad S = T$$

holds only if $(\mu^*)^{-1}(0) = \emptyset$. That is, d_{μ^*} is a metric if and only if the only set of μ^*-measure zero is the empty set. We claim that $\mu^* \colon \mathbf{2}^X \to \mathbb{R}_{\geq 0}$ is continuous with respect to the semimetric topology***missing stuff*** defined by d_{μ^*}. To see this, let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \epsilon$. Then, if S, T satisfy $d_{\mu^*}(S, T) < \delta$, we have

$$\begin{aligned} |\mu^*(S) - \mu^*(T)| &= |\mu^*(S \triangle \emptyset) - \mu^*(T \triangle \emptyset)| \\ &= |d_{\mu^*}(S, \emptyset) - d_{\mu^*}(T, \emptyset)| \leq d_{\mu^*}(S, T) = \epsilon, \end{aligned}$$

[5] Hans Hahn (1879–1934) was an Austrian mathematician whose contributions to mathematics were primarily in the areas of set theory and functional analysis. Andrey Nikolaevich Kolmogorov (1903–1987) is an important Russian mathematician. He made essential contributions to analysis, algebra, and dynamical systems. He also established the axiomatic foundations of probability theory.

using Exercise 1.1.2 and Proposition **??** (noting that this holds for semimetrics, as well as for metrics).***missing stuff***

Now define $\mathrm{cl}(\mathscr{A})$ to be the closure of $\mathscr{A} \subseteq \mathbf{2}^X$ using the semimetric d_{μ^*}. Thus $B \in \mathrm{cl}(\mathscr{A})$ if there exists a sequence $(A_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{A}$ such that $\lim_{j\to\infty} d_{\mu^*}(B, A_j) = 0$. We claim that $\mathrm{cl}(\mathscr{A})$ is a σ-algebra. Certainly $\emptyset \in \mathrm{cl}(\mathscr{A})$. Let $B \in \mathrm{cl}(\mathscr{A})$. Then there exists a sequence $(A_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{A}$ such that $\lim_{j\to\infty} d_{\mu^*}(B, A_j) = 0$. Using Exercise 1.1.2 we have

$$d_{\mu^*}(X \setminus B, X \setminus A_j) = d_{\mu^*}(B, A_j), \qquad j \in \mathbb{Z}_{>0}.$$

Thus

$$\lim_{j\to\infty} d_{\mu^*}(X \setminus B, X \setminus A_j) = 0$$

and so $X \setminus B \in \mathrm{cl}(\mathscr{A})$. Now let $B, C \in \mathrm{cl}(\mathscr{A})$ and let $(S_j)_{j\in\mathbb{Z}_{>0}}$ and $(T_j)_{j\in\mathbb{Z}_{>0}}$ be sequences in $\mathscr{A}$ such that

$$\lim_{j\to\infty} d_{\mu^*}(B, S_j) = 0, \quad \lim_{j\to\infty} d_{\mu^*}(C, T_j) = 0.$$

Then

$$\begin{aligned}\lim_{j\to\infty} d_{\mu^*}(B \cup C, S_j \cup T_j) = &\lim_{j\to\infty} \mu^*((B \cup C)\triangle(S_j \cup T_j))\\ \leq &\lim_{j\to\infty} \mu^*((B\triangle S_j) \cup (C\triangle T_j))\\ \leq &\lim_{j\to\infty} \mu^*(B\triangle S_j) + \lim_{j\to\infty} \mu^*(C\triangle T_j) = 0,\end{aligned}$$

using Exercise 1.1.2. Thus $B \cup C \in \mathrm{cl}(\mathscr{A})$. This shows that $\mathrm{cl}(\mathscr{A})$ is an algebra. Now let $(B_j)_{j\in\mathbb{Z}_{>0}}$ be a countable family of subsets from $\mathrm{cl}(\mathscr{A})$. Define $C_k = \cup_{j=1}^k B_j$ so that $C_k \in \mathrm{cl}(\mathscr{A})$, $k \in \mathbb{Z}_{>0}$. Then

$$\begin{aligned}\lim_{k\to\infty} d_{\mu^*}(\cup_{j\in\mathbb{Z}_{>0}} B_j, C_k) = &\lim_{k\to\infty} \mu^*((\cup_{j\in\mathbb{Z}_{>0}} B_j)\triangle(\cup_{j=1}^k B_j))\\ \leq &\lim_{k\to\infty} \mu^*(\cup_{j=k+1}^\infty B_j).\end{aligned}$$

Since $\mu^*(X) < \infty$ by assumption, the sequence $(\mu^*(\cup_{j=1}^k B_j))_{k\in\mathbb{Z}_{>0}}$ is a bounded monotonically increasing sequence, and so converges. This implies that

$$\lim_{k\to\infty} d_{\mu^*}(\cup_{j\in\mathbb{Z}_{>0}} B_j, C_k) = \lim_{k\to\infty} \mu^*(\cup_{j=k+1}^\infty B_j) = 0.$$

Thus $\cup_{j\in\mathbb{Z}_{>0}} B_j \in \mathrm{cl}(\mathscr{A})$ since $\mathrm{cl}(\mathscr{A})$ is closed ***missing stuff*** and since $C_k \in \mathrm{cl}(\mathscr{A})$ for each $k \in \mathbb{Z}_{>0}$. This shows that $\mathrm{cl}(\mathscr{A})$ is a σ-algebra, as desired.

We will now show that $\mu^*|\,\mathrm{cl}(\mathscr{A})$ is a measure. We certainly have $\mu^*(\emptyset) = 0$. We next claim that $\mu^*|\,\mathrm{cl}(\mathscr{A})$ is finitely-additive. To see this, let $B, C \in \mathrm{cl}(\mathscr{A})$ be disjoint and let $(S_j)_{j\in\mathbb{Z}_{>0}}$ and $(T_j)_{j\in\mathbb{Z}_{>0}}$ be sequences in $\mathscr{A}$ such that

$$\lim_{j\to\infty} d_{\mu^*}(B, S_j) = 0, \quad \lim_{j\to\infty} d_{\mu^*}(C, T_j) = 0.$$

We then have, using continuity of μ^* and additivity of $\mu^*|\mathscr{A} = \mu_0$,

$$\mu^*(B \cup C) = \lim_{j\to\infty} \mu^*(S_j \cup T_j) = \lim_{j\to\infty} \mu^*(S_j) + \lim_{j\to\infty} \mu^*(T_j - S_j) = \mu^*(B) + \mu^*(C).$$

A simple induction then gives finite-additivity. Finally, let $(B_j)_{j\in\mathbb{Z}_{>0}}$ be a countable collection of disjoint sets from $\mathrm{cl}(\mathscr{A})$. Because μ^* is an outer measure we have

$$\mu^*\Big(\bigcup_{j\in\mathbb{Z}_{>0}} B_j\Big) \leq \sum_{j=1}^{\infty} \mu^*(B_j).$$

Since $\mu^*|\,\mathrm{cl}(\mathscr{A})$ is finitely-additive we have

$$\mu^*\Big(\bigcup_{j\in\mathbb{Z}_{>0}} B_j\Big) \geq \mu^*\Big(\bigcup_{j=1}^{k} B_j\Big) = \sum_{j=1}^{k} \mu^*(B_j)$$

for every $k \in \mathbb{Z}_{>0}$. Thus

$$\mu^*\Big(\bigcup_{j\in\mathbb{Z}_{>0}} B_j\Big) \geq \sum_{j=1}^{\infty} \mu^*(B_j),$$

which allows us to conclude countable-additivity of $\mu^*|\,\mathrm{cl}(\mathscr{A})$.

Since $\mathscr{A} \subseteq \mathrm{cl}(\mathscr{A})$ it follows from Proposition 5.2.7 that $\sigma(\mathscr{A}) \subseteq \mathrm{cl}(\mathscr{A})$. Since $\mu^*|\,\mathrm{cl}(\mathscr{A})$ is a measure, it is surely also true that $\mu \triangleq \mu^*|\sigma(\mathscr{A})$ is a measure. This proves the existence assertion of the theorem under the assumption that $\mu_0(X) < \infty$.

For uniqueness, let $\tilde{\mu}\colon \mathrm{cl}(\mathscr{A}) \to \mathbb{R}_{\geq 0}$ be a measure having the property that $\tilde{\mu}|\mathscr{A} = \mu_0$. Let $B \in \sigma(\mathscr{A})$ and let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a family of subsets such that $B \subseteq \cup_{j\in\mathbb{Z}_{>0}} A_j$. Since $\tilde{\mu}|\mathscr{A} = \mu_0$ we have

$$\tilde{\mu}(B) \leq \sum_{j=1}^{\infty} \tilde{\mu}(A_j) = \sum_{j=1}^{\infty} \mu_0(A_j),$$

using Proposition 5.3.10. From this we infer that

$$\tilde{\mu}(B) \leq \inf\Big\{\sum_{j=1}^{\infty} \mu_0(A_j) \;\Big|\; B \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} A_j,\ A_j \in \mathscr{A},\ j \in \mathbb{Z}_{>0}\Big\} = \mu(B).$$

In like manner we have that $\tilde{\mu}(X \setminus B) \leq \mu(X \setminus B)$. Thus

$$\tilde{\mu}(B) = \tilde{\mu}(X) - \tilde{\mu}(X \setminus B) \geq \mu(X) - \mu(X \setminus B) = \mu(B).$$

Thus $\tilde{\mu}(B) = \mu(B)$, as desired.

Finally, we prove the theorem, removing the assumption that $\mu_0(X) < \infty$. Since the hypotheses of the theorem include μ_0 being σ-finite, there exists a countable collection $(Y_j)_{j\in\mathbb{Z}_{>0}}$ of subsets from $\mathscr{A}$ such that $\mu_0(Y_j) < \infty$, $j \in \mathbb{Z}_{>0}$, and such that $X = \cup_{j\in\mathbb{Z}_{>0}} Y_j$. Then define

$$\begin{aligned}
X_1 &= Y_1\\
X_2 &= Y_2 \cap (X \setminus Y_1)\\
&\;\vdots\\
X_j &= Y_j \cap (X \setminus Y_1) \cap \cdots \cap (X \setminus Y_{j-1})\\
&\;\vdots
\end{aligned}$$

noting that the family of sets $(X_j)_{j\in\mathbb{Z}_{>0}}$ is in $\mathscr{A}$ since $\mathscr{A}$ is an algebra. Moreover, by construction the sets $(X_j)_{j\in\mathbb{Z}_{>0}}$ are pairwise disjoint, have the property that $\mu(X_j) < \infty$, $j \in \mathbb{Z}_{>0}$, and satisfy $X = \cup_{j\in\mathbb{Z}_{>0}} X_j$. Denote

$$\mathscr{A}_j = \{X_j \cap A \mid A \in \mathscr{A}\}, \quad \sigma(\mathscr{A})_j = \{X_j \cap B \mid B \in \sigma(\mathscr{A})\}, \quad \mu_{0,j} = \mu_0|\mathscr{A}_j.$$

We claim that $\sigma(\mathscr{A})_j = \sigma(\mathscr{A}_j)$. To show this one must show that $\sigma(\mathscr{A})_j$ is a σ-algebra on X_j containing $\mathscr{A}_j$ and that any σ-algebra containing on X_j containing $\mathscr{A}_j$ contains $\sigma(\mathscr{A})_j$. It is a straightforward exercise manipulating sets to show that $\sigma(\mathscr{A})_j$ is a σ-algebra containing $\mathscr{A}_j$, and we leave this to a sufficiently bored reader. So let $\mathscr{A}'_j$ be a σ-algebra on X_j containing $\mathscr{A}_j$. Let

$$\mathscr{A}' = \{A \cup B \mid A \in \mathscr{A}'_j, B = (X \setminus X_j) \cap B',\ B' \in \sigma(\mathscr{A})\}.$$

By Exercise 5.2.5 we conclude that $\mathscr{A}'$ is a σ-algebra on $X = X_j \cup (X \setminus X_j)$. Moreover, $\mathscr{A} \subseteq \mathscr{A}'$ and so $\sigma(\mathscr{A}) \subseteq \mathscr{A}'$. But this means that if $X_j \cap B \in \sigma(\mathscr{A})_j$ then $X_j \cap B \in \mathscr{A}'_j$, giving our claim.

Now note that, for each $j \in \mathbb{Z}_{>0}$, the data X_j, $\mathscr{A}_j$, and $\mu_{0,j}$ satisfy the hypotheses used in the first part of the proof. Therefore, there exists a measure μ_j on $\sigma(\mathscr{A}_j) = \sigma(\mathscr{A})_j$ agreeing with $\mu_{0,j}$ on $\mathscr{A}_j$. Now define $\mu\colon \sigma(\mathscr{A}) \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\mu(B) = \sum_{j=1}^{\infty} \mu_j(X_j \cap B).$$

That μ is a measure is easily verified using the fact that μ_j, $j \in \mathbb{Z}_{>0}$ is a measure and that the family of sets $(X_j)_{j\in\mathbb{Z}_{>0}}$ is pairwise disjoint. We leave the straightforward working out of this to the, again sufficiently bored, reader. It is also clear that $\mu|\mathscr{A} = \mu_0$. This gives the existence part of the proof. For uniqueness, suppose that $\tilde{\mu}\colon \sigma(\mathscr{A}) \to \overline{\mathbb{R}}_{\geq 0}$ is a measure such that $\tilde{\mu}|\mathscr{A} = \mu_0$ and let $B \in \sigma(\mathscr{A})$. By uniqueness from the first part of the proof we have $\tilde{\mu}(X_j \cap B) = \mu(X_j \cap B)$. Therefore, by countable-additivity of $\tilde{\mu}$,

$$\tilde{\mu}(B) = \sum_{j=1}^{\infty} \tilde{\mu}(X_j \cap B) = \sum_{j=1}^{\infty} \mu(X_j \cap B) = \mu(B),$$

as desired. ■

The proof of the preceding theorem introduced an important construction. As we shall not make use of this in any subsequent part of the text, let us expound a little on this here.

5.3.27 Remark (Semimetrics and measures) A key ingredient in our proof of the Hahn–Kolmogorov Extension Theorem was a semimetric associated with a measure. This construction can be generalised somewhat. Let X be a set, let $\mathscr{S} \subseteq \mathbf{2}^X$, and let $\mu\colon \mathscr{S} \to \mathbb{R}_{\geq 0}$ be a finite-valued finitely-subadditive measure, i.e.,

$$\mu\Big(\cup_{j=1}^{k} A_j\Big) \leq \sum_{j=1}^{k} \mu(A_j), \qquad A_1, \dots, A_k \in \mathscr{A}.$$

Then we define $d_\mu\colon \mathscr{S}\times\mathscr{S}\to\mathbb{R}_{\geq 0}$ by $d_\mu(S,T)=\mu(S\triangle T)$, recalling from Section 1.1.2 the definition of the symmetric complement $\triangle$. As in the above proof, we can verify that d_μ is a semimetric, and is a metric if and only if the only set of measure zero is the empty set. If μ is not finite-valued, then we can instead use

$$d'_\mu(S,T)=\max\{1,\mu(S\triangle T)\},$$

with the same conclusions.

In the proof we used this semimetric to define, in a topological sense, the closure $\mathrm{cl}(\mathscr{A})$ of the algebra $\mathscr{A}$, and we showed that $\sigma(\mathscr{A})\subseteq\mathrm{cl}(\mathscr{A})$. In fact, although we did not need this in the proof above, $\mathrm{cl}(\mathscr{A})$ is the completion of $\sigma(\mathscr{A})$. This gives a neat loop-closing for the use of the word "completion" in this context, since it gives this a standard topological meaning. The Hahn–Kolmogorov Extension Theorem, then, becomes sort of a result about the extension of uniformly continuous functions to the completion, *a la* Theorem **??**. When one digs more deeply into measure theory *per se*, these sorts of matters become more important. •

Now let us both specialise and extend our discussion of outer measures generated by a collection of subsets. We consider in detail the situation where we begin with a measure space.

5.3.28 Definition (Inner and outer measure of a measure) Let $(X,\mathscr{A},\mu)$ be a measure space.

(i) The ***outer measure*** associated to μ is the map $\mu^*\colon \mathbf{2}^X\to\overline{\mathbb{R}}_{\geq 0}$ defined by

$$\mu^*(S)=\inf\{\mu(A)\mid A\in\mathscr{A},\ S\subseteq A\}.$$

(ii) The ***inner measure*** associated to μ is the map $\mu_*\colon \mathbf{2}^X\to\overline{\mathbb{R}}_{\geq 0}$ defined by

$$\mu_*(S)=\sup\{\mu(A)\mid A\in\mathscr{A},\ A\subseteq S\}.$$ •

The following corollary to Proposition 5.3.24 answers one of the basic questions we raised upon defining the concept of an outer measure.

5.3.29 Corollary (The outer measure of a measure is an outer measure) *If $(X,\mathscr{A},\mu)$ be a measure space then the outer measure μ^* associated to μ is an outer measure as per Definition 5.3.4.*

Proof Since a σ-algebra is an algebra and since countable unions of measurable sets are measurable, this follows directly from Proposition 5.3.24. ■

One way to interpret the preceding result is that it provides a natural way of extending a measure, possibly only defined on a strict subset of the collection of all subsets, to a means of measuring "size" for all subsets, and that this extension is, in fact, an outer measure. This provides, then, a nice characterisation how a measure approximates sets "from above." What about the rôle of the inner measure that approximates sets "from below"? The following result clarifies this rôle, and illustrates one place where completeness is important.

5.3.30 Proposition (Sets for which inner and outer measure agree are in the completion) *Let* $(\mathrm{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathrm{A} \subseteq \mathrm{X}$ *be such that* $\mu^*(\mathrm{A}) < \infty$*. Then* $\mu_*(\mathrm{A}) = \mu^*(\mathrm{A})$ *if and only if* $\mathrm{A} \in \mathscr{A}_\mu$*.*

Proof Suppose that $A \in \mathscr{A}_\mu$ and let $L, U \in \mathscr{A}$ satisfy $L \subseteq A \subseteq U$ and $\mu(U \setminus L) = 0$. Then

$$\mu(L) \leq \mu_*(A) \leq \mu^*(A) \leq \mu(U),$$

giving $\mu_*(A) = \mu^*(A)$ since $\mu(L) = \mu(U)$.

Conversely, suppose that $\mu_*(A) = \mu^*(A)$. Let $k \in \mathbb{Z}_{>0}$. Then there exists sets $M_k, V_k \in \mathscr{A}$ such that $M_k \subseteq A \subseteq V_k$ and such that

$$\mu_*(A) < \mu(M_k) + \tfrac{1}{k}, \quad \mu(V_k) < \mu^*(A) + \tfrac{1}{k}.$$

Then, for $k \in \mathbb{Z}_{>0}$ define

$$L_k = \cup_{j=1}^k M_j \in \mathscr{A}, \quad U_k = \cap_{j=1}^k V_j \in \mathscr{A},$$

noting that $M_k \subseteq L_k \subseteq A$, $A \subseteq U_k \subseteq V_k$, $L_k \subseteq L_{k+1}$, and $U_{k+1} \subseteq U_k$ for $k \in \mathbb{Z}_{>0}$. We then have

$$\mu_*(A) - \tfrac{1}{k} < \mu(M_k) \leq \mu(L_k) \leq \mu(U_k) \leq \mu(V_k) < \mu^*(A) + \tfrac{1}{k}.$$

Taking the limit as $k \to \infty$ gives

$$\lim_{k\to\infty} \mu(L_k) = \lim_{k\to\infty} \mu(L_k).$$

If we define $L = \cup_{k\in\mathbb{Z}_{>0}} L_k \in \mathscr{A}$ and $U = \cap_{k\in\mathbb{Z}_{>0}} U_k \in \mathscr{A}$ then we have $L \subseteq A \subseteq U$ and, by Proposition 5.3.10, $\mu(L) = \mu(U)$. Thus $A \subseteq \mathscr{A}_\mu$. ■

5.3.5 Probability measures

In this section we introduce the notion of a probability measure. As the name suggests, probability measures arise naturally in the study of probability theory, but this is something we will not take up here, postponing a general study of this for ***missing stuff***.

Let us first define what we mean by a probability measure.

5.3.31 Definition (Probability space, probability measure) A ***probability space*** is a measure space $(X, \mathscr{A}, \mu)$ for which $\mu(X) = 1$. The set X is called the ***sample space***, the σ-algebra $\mathscr{A}$ is called the set of ***events***, and the measure μ is called a ***probability*** measure. •

Let us give some examples.

5.3.32 Examples (Probability spaces)

1. Let us consider the classical example of a problem in so-called "discrete probability." We suppose that we have a coin which, when we flip it, has two outcomes, denoted "H" for "heads" and "T" for "tails." Let us suppose that we know that the coin is biased in a known way, so that the likelihood of seeing a head on any flip is $p \in [0, 1]$. Then the likelihood of seeing a tail on any flip is $1 - p$. We shall flip this coin once, and the record the outcome. Thus the

sample space is $X = \{\mathrm{H}, \mathrm{T}\}$. The σ-algebra of events we take to be $\mathscr{A} = \mathbf{2}^X$. Thus there are four events: (a) $\emptyset$ (corresponding to an outcome of neither "heads" nor "tails"); (b) $\{\mathrm{H}\}$ (corresponding to an outcome of "heads"); (c) $\{\mathrm{T}\}$ (corresponding to an outcome of "tails"); (d) $\{\mathrm{H}, \mathrm{T}\}$ (corresponding to an outcome of either "heads" or "tails"). The probability measure is defined by

$$\mu(\{\mathrm{H}\}) = p, \quad \mu(\{\mathrm{T}\}) = (1-p).$$

The probability measure for the events $\emptyset$ and $\{\mathrm{H}, \mathrm{T}\}$ must be 0 (because the measure of the empty set is always zero) and 1 (by countable additivity of the measure), respectively. Thus μ is a probability measure.

2. We have a biased coin as above. But now we perform an trial where we flip the coin n times and record the outcome each time. An element of the sample space X is an outcome of a single trial. Thus an element of the sample space is an element of $X = \{\mathrm{H}, \mathrm{T}\}^{\{1,\dots,n\}}$, the set of maps from $\{1, \dots, n\}$ to $\{\mathrm{H}, \mathrm{T}\}$. Note that $\operatorname{card}(X) = 2^n$. If $\phi \in X$ then the outcome of this trial is represented by the sequence

$$(\phi(1), \dots, \phi(n)) \in \{\mathrm{H}, \mathrm{T}\}^n.$$

The σ-algebra defining the set of events is the set of subsets of all trials: $\mathscr{A} = \mathbf{2}^X$. Now let us define a meaningful probability measure. For a trial $\phi \in X$ let $n_{\mathrm{H}}(\phi)$ be the number of heads appearing in the trial and let $n_{\mathrm{T}}(\phi)$ be the number of tails appearing in the trial. Obviously, $n_{\mathrm{H}}(\phi) + n_{\mathrm{T}}(\phi) = n$ for every $\phi \in X$. We then define

$$\mu(\phi) = p^{n_{\mathrm{H}}(\phi)}(1-p)^{n_{\mathrm{T}}(\phi)}.$$

This then defines μ on $\mathbf{2}^X$ by countable additivity. We should check that this is a probability measure, i.e., that $\mu(X) = 1$. For fixed $k \in \{1, \dots, k\}$, the number of trials in which k heads appears is

$$\binom{n}{k} \triangleq \frac{n!}{k!(n-k)!},$$

i.e., the binomial coefficient $B_{n,k}$ from Exercise 2.2.1. Note that, according to Exercise 2.2.1,

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

Since the expression on the left is the sum over the trials with any possible number of heads, it is the sum over all possible trials.

3. Consider the problem of "randomly" choosing a number in the interval $[0, 1]$. Thus $X = [0, 1]$. We wish to use the Lebesgue measure as a probability measure. Note that, according to our constructions of Section 5.4, to do this pretty much necessitates taking $\mathscr{A} = \mathscr{L}([0, 1])$ as the set of events.

4. Let $x_0 \in \mathbb{R}$ and let $\sigma \in \mathbb{R}_{>0}$. Let us consider the sample space $X = \mathbb{R}$, the set of events $\mathscr{A} = \mathscr{L}(\mathbb{R})$, and the measure $\gamma_{x_0,\sigma} \colon \mathscr{L}(\mathbb{R}) \to \mathbb{R}$ defined by

$$\gamma_{x_0,\sigma}(A) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} \chi_A(x) \exp(-\tfrac{1}{2\sigma^2}(x - x_0)^2) \, \mathrm{d}x.$$

We claim that $\gamma_{x_0,\sigma}$ is a probability measure, i.e., that $\gamma_{x_0,\sigma}(\mathbb{R}) = 1$. The following lemma is useful in verifying this.

1 Lemma $\int_{\mathbb{R}} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi = \sqrt{\pi}.$

Proof By Fubini's Theorem we write

$$\Big(\int_{\mathbb{R}} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi\Big)^2 = \Big(\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x\Big)\Big(\int_{\mathbb{R}} \mathrm{e}^{-y^2}\,\mathrm{d}y\Big) = \int_{\mathbb{R}^2} \mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\mathrm{d}y.$$

By Example **??–??** we have

$$\Big(\int_{\mathbb{R}} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi\Big)^2 = \int_{\mathbb{R}_{>0}\times[-\pi,\pi]} r\mathrm{e}^{-r^2}\mathrm{d}r\mathrm{d}\theta = 2\pi\int_{\mathbb{R}_{>0}} r\mathrm{e}^{-r^2}\,\mathrm{d}r.$$

Now we make another change of variable $\rho = r^2$ to obtain

$$\Big(\int_{\mathbb{R}} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi\Big)^2 = \pi\int_{\mathbb{R}_{>0}} \mathrm{e}^{-\rho}\,\mathrm{d}\rho = \pi,$$

and so we get the result. ▼

By making the change of variable $\xi = \frac{1}{\sqrt{2}\sigma}(x - x_0)$, we can then directly verify that $\gamma_{x_0,\sigma}(\mathbb{R}) = 1$. This probability measure is called the ***Gaussian measure*** with ***mean*** x_0 and ***variance*** σ. •

5.3.6 Product measures

In Section 5.2.3 we showed how algebras on the factors of a product give algebras and σ-algebras on the product. In this section we investigate how to define measures on products given measures on each of the factors. The procedure for this is surprisingly technical; we use the Hahn–Kolmogorov Extension Theorem. It is also possible to define measures on products using the integral, after the integral has been defined. We refer to Section 5.8.1 for this construction.

For now, let us state and prove the basic result concerning the construction of measures on products of measure spaces.

5.3.33 Theorem (Measures on products of measure spaces) *If* $(\mathsf{X}_j, \mathscr{A}_j, \mu_j)$, $j \in \{1, \dots, k\}$, *are σ-finite measure spaces then there exists a unique measure*

$$\mu_1 \times \cdots \times \mu_k \colon \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k) \to \overline{\mathbb{R}}_{\geq 0}$$

such that

$$\mu_1 \times \cdots \times \mu_k(\mathsf{A}_1 \times \cdots \times \mathsf{A}_k) = \mu_1(\mathsf{A}_1)\cdots\mu_k(\mathsf{A}_k)$$

for every $\mathsf{A}_1 \times \cdots \times \mathsf{A}_k \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_k$.

Proof We use a couple of technical lemmata.

1 Lemma *Let* X *be a set and let* $\mathscr{S}_0 \subseteq \mathbf{2}^X$ *be a family of subsets for which*

(i) $S_1 \cap S_2 \in \mathscr{S}_0$ *for every* $S_1, S_2 \in \mathscr{S}_0$ *and*

(ii) if $S \in \mathscr{S}_0$ *then* $X \setminus S = S_1 \cup \cdots \cup S_k$ *for some pairwise disjoint* $S_1, \dots, S_k \in \mathscr{S}_0$.

Then $\sigma_0(\mathscr{S}_0)$ *is equal to the collection of finite unions of sets from* $\mathscr{S}_0$ *and, if* $\mu_0 \colon \mathscr{S}_0 \to \overline{\mathbb{R}}_{\geq 0}$ *is finitely-additive, then there exists a unique finitely-additive function* $\mu_0 \colon \sigma_0(\mathscr{S}) \to \overline{\mathbb{R}}_{\geq 0}$ *such that* $\mu|\mathscr{S}_0 = \mu_0$.

Proof First we claim that the set of finite unions of sets from $\mathscr{S}_0$, let us denote this collection of subsets by $\overline{\mathscr{S}}_0$, is an algebra. To see that $X \in \overline{\mathscr{S}}_0$, let $S \in \mathscr{S}_0$ and write, by hypothesis,

$$X = S \cup (X \setminus S) = S \cup (S_1 \cdots \cup S_k)$$

for some $S_1, \dots, S_k \in \mathscr{S}_0$. Thus $X \in \overline{\mathscr{S}}_0$. Now let $S \in \overline{\mathscr{S}}_0$ and write $S = S_1 \cup \cdots \cup S_k$ for $S_1, \dots, S_k \in \mathscr{S}_0$. Then, by De Morgan's Laws,

$$X \setminus S = (X \setminus S_1) \cap \cdots \cap (X \setminus S_k).$$

Thus $X \setminus S$ is, by assumption, a finite intersection of finite unions of sets from $\mathscr{S}_0$. Since intersections of finitely many sets from $\mathscr{S}_0$ are in $\mathscr{S}_0$, it then follows that $X \setminus S \in \overline{\mathscr{S}}_0$. Thus, by Exercise 5.2.1, $\overline{\mathscr{S}}_0$ is an algebra. Moreover, if $\mathscr{A}$ is any algebra containing $\mathscr{S}_0$ then $\mathscr{A}$ must necessarily contain the finite unions of sets from $\mathscr{S}_0$. Thus $\overline{\mathscr{S}}_0 \subseteq \mathscr{A}$. By Proposition 5.2.8 this shows that $\overline{\mathscr{S}}_0 = \sigma_0(\mathscr{S}_0)$, as desired.

Now let $A \in \sigma_0(\mathscr{S}_0)$ so that $A = A_1 \cup \cdots \cup A_k$ for some $A_1, \dots, A_k \in \mathscr{S}_0$. By Lemma 1 in the proof of Proposition 5.3.2, there are then *disjoint* sets $T_1, \dots, T_m \in \mathscr{S}_0$ such that $A = T_1 \cup \cdots \cup T_m$. We then define

$$\mu(A) = \mu_0(T_1) + \cdots + \mu_0(T_m).$$

We must show that this definition is independent of the particular way in which one writes A as a disjoint union of sets from $\mathscr{S}_0$. Suppose that $A = T'_1 \cup \cdots \cup T'_n$ for disjoint $T'_1, \dots, T'_n \in \mathscr{S}_0$. Then

$$A = \cup_{j=1}^m T_j = \cup_{l=1}^n T'_l = \cup_{j=1}^m \cup_{l=1}^n T_j \cap T'_l,$$

as may be easily verified. It then follows that

$$\mu\big(\cup_{j=1}^m T_j\big) = \sum_{j=1}^m \mu_0(T_j) = \sum_{j=1}^m \sum_{l=1}^n \mu_0(T_j \cap T'_l) = \sum_{l=1}^n \sum_{j=1}^m \mu_0(T'_l \cap T_j) = \sum_{l=1}^n \mu_0(T'_l),$$

giving the well-definedness of μ, and so the existence assertion of the lemma. Uniqueness follows immediately from finite-additivity of μ. ▼

2 Lemma *For sets* $X_1, \dots, X_k$ *with algebras* $\mathscr{A}_j \subseteq \mathbf{2}^{X_j}$, $j \in \{1, \dots, k\}$, *let* $\mu_j \colon \mathscr{A}_j \to \overline{\mathbb{R}}_{\geq 0}$, $j \in \{1, \dots, k\}$, *be finitely-additive. Then there exists a unique finitely-additive*

$$\mu \colon \sigma_0(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k) \to \overline{\mathbb{R}}_{\geq 0}$$

such that

$$\mu(A_1 \times \cdots \times A_k) = \mu_1(A_1) \cdots \mu_k(A_k) \tag{5.4}$$

for every $A_j \in \mathscr{A}_j$, $j \in \{1, \dots, k\}$.

Proof Let us abbreviate $\mathscr{A} = \sigma_0(\mathscr{A}_1 \times \cdots \times \mathscr{A}_k)$. By Proposition 5.2.16, if $A \in \mathscr{A}$ then we can write

$$A = R_1 \cup \cdots \cup R_m$$

for disjoint measurable rectangles $R_1, \dots, R_m$. We then define

$$\mu(A) = \mu(R_1) + \cdots + \mu(R_m), \tag{5.5}$$

where $\mu(R_j)$, $j \in \{1, \dots, m\}$, is defined as in (5.4). We must show that this definition of μ is independent of the way in which one expresses A as a finite disjoint union of measurable rectangles. First let us suppose that

$$A = A_1 \times \cdots \times A_k \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_k.$$

We shall prove by induction on k that if A is written as a finite disjoint union of measurable rectangles, $A = R_1 \cup \cdots \cup R_m$, that (5.5) holds. This assertion is vacuous for $k = 1$, so assume it holds for $k = n - 1$ and let

$$A_1 \times \cdots \times A_n = \cup_{j=1}^m B'_j \times B_j$$

where $B'_j \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_{n-1}$ and $B_j \in \mathscr{A}_n$ for each $j \in \{1, \dots, m\}$. By the induction hypothesis and by our knowing the volumes of measurable rectangles, there exists a finitely-additive function $\mu' \colon \sigma_0(\mathscr{A}_1 \times \cdots \times \mathscr{A}_{n-1}) \to \overline{\mathbb{R}}_{\geq 0}$ such that

$$\mu'(A'_1 \times \cdots \times A'_{n-1}) = \mu_1(A'_1) \cdots \mu_{n-1}(A'_{n-1})$$

for every $A'_1 \times \cdots \times A'_{n-1} \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_{n-1}$. We are charged with showing that

$$\begin{aligned}\mu(A_1 \times \cdots \times A_n) &= \mu_1(A_1) \cdots \mu_{n-1}(A_{n-1})\mu_n(A_n) \\ &= \mu'(A_1 \times \cdots \times A_{n-1})\mu_n(A_n) = \sum_{j=1}^m \mu'(B'_j)\mu_n(B_j),\end{aligned}$$

the last equality being the only that is not obvious.

From Lemma 1 in the proof of Proposition 5.3.2, there exists pairwise disjoint sets $C_1, \dots, C_r \subseteq A_n$ such that each of the sets $B_1, \dots, B_k$ is a finite union of the sets $C_1, \dots, C_r$. Thus, for each $j \in \{1, \dots, k\}$, there exists pairwise disjoint sets $S_{j1}, \dots, S_{jm_j} \subseteq A_n$, taken from the collection of sets $C_1, \dots, C_r$, for which $B_j = S_{j1} \cup \cdots \cup S_{jm_j}$. Thus

$$A_1 \times \cdots \times A_n = \cup_{j=1}^m B'_j \times \left(\cup_{l_j=1}^{m_j} S_{jl_j}\right) = \cup_{j=1}^m \cup_{l_j=1}^{m_j} B'_j \times S_{jl_j}.$$

Now, for each $s \in \{1, \dots, r\}$, let $J_s \subseteq \{1, \dots, k\}$ be defined so that $j \in J_s$ if and only if there exists $l_j \in \{1, \dots, m_j\}$ (necessarily unique) such that $S_{jl_j} = C_s$. Then define $B''_s = \cup_{j \in J_s} B'_j$. Since the measurable rectangles $B'_j \times B_j$, $j \in \{1, \dots, k\}$, are pairwise disjoint, it follows that the measurable rectangles B'_j, $j \in J_s$, are pairwise disjoint. Also note that we then have

$$A_1 \times \cdots \times A_n = \cup_{s=1}^r B''_s \times C_s,$$

noting that $C_1, \dots, C_s$ are pairwise disjoint. This implies that $\cup_{s=1}^r C_s = A_n$. This, in turn, forces us to conclude that $B''_s = A_1 \times \cdots \times A_{n-1}$ for each $s \in \{1, \dots, r\}$.

Now let us use the above facts, along with the induction hypothesis. Finite-additivity of μ_n gives

$$\mu_n(B_j) = \sum_{l_j=1}^{m_j} \mu_n(S_{jl_j}), \qquad j \in \{1, \dots, k\},$$

and

$$\sum_{s=1}^{r} \mu_n(C_s) = \mu_n(A_n).$$

Also, finite-additivity of μ' gives

$$\mu'(A_1 \times \cdots \times A_{n-1}) = \mu'(\cup_{j \in J_s} B'_j) = \sum_{j \in J_s} \mu'(B_j).$$

Putting this all together gives

$$\begin{aligned}\sum_{j=1}^{k} \mu'(B'_j)\mu_n(B_j) = \sum_{j=1}^{k} \mu'(B_n) \sum_{l_j=1}^{m_j} \mu_n(S_{jl_j}) &= \sum_{s=1}^{r} \sum_{j \in J_s} \mu'(B'_j)\mu_n(C_s) \\ &= \mu'(A_1 \times \cdots \times A_{n-1})\mu_n(A_n).\end{aligned}$$

This proves that the definition of volume of measurable rectangles is independent of how these rectangles are decomposed into finite disjoint unions of measurable rectangles.

The existence part of the lemma now follows from Lemma 1, along with Proposition 5.2.16. Uniqueness immediately follows from Proposition 5.2.16, along with the uniqueness assertion from Lemma 1. ▼

We complete the proof by induction on k, the assertion being clear when $k = 1$. So suppose that the conclusions of the theorem hold for $k = 1, \dots, m-1$ for some $m \geq 2$, and let $(X_j, \mathscr{A}_j, \mu_j)$, $j \in \{1, \dots, m\}$, be measure spaces satisfying the hypotheses of the theorem. Let us denote $Y = X_1 \times \cdots \times X_{m-1}$ and $Z = X_m$, $\mathscr{B} = \sigma(\mathscr{A}_1 \times \cdots \times \mathscr{A}_{m-1})$ and $\mathscr{C} = \mathscr{A}_m$, and $\nu = \mu_1 \times \cdots \times \mu_{m-1}$ and $\lambda = \mu_m$. We use the induction hypothesis to define $\nu\colon \mathscr{B} \to \overline{\mathbb{R}}_{\geq 0}$. We now wish to show that there exists a unique map $\nu \times \lambda\colon \mathscr{B} \times \mathscr{C} \to \overline{\mathbb{R}}_{\geq 0}$ such that

$$\nu \times \lambda(B \times C) = \nu(B)\lambda(C)$$

for every $B \times C \in \mathscr{B} \times \mathscr{C}$. Note that by Proposition 5.2.16 and Lemma 2, and since a countably-additive measure is also finitely-additive, there exists a unique finitely-additive measure $\nu_0\colon \sigma_0(\mathscr{B} \times \mathscr{C}) \to \overline{\mathbb{R}}_{\geq 0}$ such that

$$\nu_0(B \times C) = \nu(B)\lambda(C)$$

for every $B \times C \in \mathscr{B} \times \mathscr{C}$. By the Hahn–Kolmogorov Extension Theorem we need only show that ν_0 is countably-additive.

Let us first suppose that ν and λ are finite. Then, by Proposition 5.3.3, it suffices to show that if $(A_j)_{j \in \mathbb{Z}_{>0}}$ is a sequence of subsets from $\sigma_0(\mathscr{B} \times \mathscr{C})$ such that $A_j \supseteq A_{j+1}$ and such that $\cap_{j \in \mathbb{Z}_{>0}} A_j = \emptyset$, then $\lim_{j \to \infty} \nu_0(A_j) = 0$. By Proposition 5.2.16, for each $j \in \mathbb{Z}_{>0}$ we have

$$A_j = \cup_{k=1}^{m_j} B_{jk} \times C_{jk}$$

for nonempty sets $B_{j1},\dots,B_{1m_j} \in \mathscr{B}$ and $C_{j1},\dots,C_{jm_j} \in \mathscr{C}$. Moreover, as we argued in the proof of Lemma 2, we may suppose without loss of generality that the sets $B_{j1},\dots,B_{jm_j}$ are pairwise disjoint. Now define $f_j\colon Y \to \mathbb{R}_{\geq 0}$ by

$$f_j(y) = \begin{cases} \lambda(C_{jk}), & y \in B_{jk}, \\ 0, & y \notin \cup_{k=1}^{m_j} B_{jk}. \end{cases}$$

For $y \in Y$ and $j \in \mathbb{Z}_{>0}$ there exists a unique $k(j,y) \in \{1,\dots,m_j\}$ such that $y \in B_{jk(j,y)}$. Moreover, if $j_1 < j_2$ we have

$$C_{j_1k(j_1,y)} = \{z \in Z \mid (y,z) \in A_{j_1}\} \subseteq \{z \in Z \mid (y,z) \in A_{j_2}\} = C_{j_2k(j_2,y)}$$

Therefore, the sequence $(f_j(y))_{j\in\mathbb{Z}_{>0}}$ is monotonically decreasing for each $y \in Y$. Moreover, $\lim_{j\to\infty} f_j(y) = 0$ since

$$\cap_{j\in\mathbb{Z}_{>0}} C_{jk(j,y)} \subseteq \cap_{j\in\mathbb{Z}_{>0}} \{z \in Z \mid (y,z) \in A_j\} = \emptyset.$$

Now let $\epsilon \in \mathbb{R}_{>0}$ and $j \in \mathbb{Z}_{>0}$ and define

$$B_{j,\epsilon} = \{y \in Y \mid f_j(y) > \epsilon\}.$$

We can easily see that $B_{j,\epsilon} \subseteq \cup_{k=1}^{m_j} B_{jk}$, that $B_{j,\epsilon} \supseteq B_{j+1,\epsilon}$ for $j \in \mathbb{Z}_{>0}$, and that $\cap_{j\in\mathbb{Z}_{>0}} B_{j,\epsilon} = \emptyset$. We therefore compute

$$\nu_0(A_j) = \sum_{k=1}^{m_j} \nu(B_{jk})\lambda(C_{jk}) \leq \nu(B_{j,\epsilon})\lambda(Z) + \nu(Y)\epsilon.$$

Since $\lim_{j\to\infty} \nu(B_{j,\epsilon}) = 0$ by Proposition 5.3.3, it follows that

$$\lim_{j\to\infty} \nu_0(A_j) \leq \epsilon\nu(Y),$$

giving $\lim_{j\to\infty} \nu_0(A_j) = 0$ since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary. This shows that ν_0 is a measure on $\sigma_0(\mathscr{B} \times \mathscr{C})$.

Next suppose that ν and λ are not finite, but are σ-finite. Then let $(S_k)_{k\in\mathbb{Z}_{>0}}$ and $(T_k)_{k\in\mathbb{Z}_{>0}}$ be subsets of Y and Z, respectively, such that $\nu(S_k) < \infty$ and $\lambda(T_k) < \infty$ for $k \in \mathbb{Z}_{>0}$, and such that $Y = \cup_{k\in\mathbb{Z}_{>0}} S_k$ and $Z = \cup_{k\in\mathbb{Z}_{>0}} T_k$. We may without loss of generality suppose that $S_k \subseteq S_{k+1}$ and $T_k \subseteq T_{k+1}$ for $k \in \mathbb{Z}_{>0}$. Let us denote

$$\mathscr{B}_k = \{B \cap S_k \mid B \in \mathscr{B}\}, \quad \mathscr{C}_k = \{C \cap T_k \mid C \in \mathscr{C}\}$$

and $\nu_k = \nu_0|sB_k \times \mathscr{C}_k$, noting from what we have already proved that ν_k is a measure. Then, for disjoint sets $(A_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{B} \times \mathscr{C}$ we have

$$\begin{aligned} \sum_{j=1}^{\infty} \nu_0(A_j) &= \sum_{j=1}^{\infty} \lim_{k\to\infty} \nu_0(A_j \cap (S_k \times T_k)) = \lim_{k\to\infty} \sum_{j=1}^{\infty} \nu_k(A_j \cap (S_k \times T_k)) \\ &= \lim_{k\to\infty} \nu_k(\cup_{j\in\mathbb{Z}_{>0}} A_j \cap (S_k \times T_k)) = \nu_0(\cup_{j\in\mathbb{Z}_{>0}} A_j). \end{aligned}$$

This shows that ν_0 is a measure on $\mathscr{B} \times \mathscr{C}$.

Finally, to complete the proof by induction, one needs only to reinstate the definitions $Y = X_1 \times \dots \times X_{m-1}$ and $Z = X_m$, $\mathscr{B} = \sigma(\mathscr{A}_1 \times \dots \times \mathscr{A}_{m-1})$ and $\mathscr{C} = \mathscr{A}_m$, and $\nu = \mu_1 \times \dots \times \mu_{m-1}$ and $\lambda = \mu_m$, and then apply the induction hypothesis. ■

Let us name the measure from the preceding theorem.

5.3.34 Definition (Product measure) If $(X_j, \mathscr{A}_j, \mu_j)$, $j \in \{1, \dots, k\}$, are σ-finite measure spaces then the measure $\mu_1 \times \cdots \times \mu_k$ is the ***product measure***. •

Let us give simple examples of product measures.

5.3.35 Examples (Product measures)

1. Let X and Y be sets with $\mathscr{A}$ and $\mathscr{B}$ σ-algebras on X and Y, respectively. Define $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ and $\nu\colon \mathscr{B} \to \overline{\mathbb{R}}_{\geq 0}$ by
$$\mu(A) = \begin{cases} 0, & A = \emptyset, \\ \infty, & A \neq \emptyset, \end{cases} \qquad \nu(B) = \begin{cases} 0, & B = \emptyset, \\ \infty, & B \neq \emptyset. \end{cases}$$
Then the map $\mu \times \nu\colon \sigma(\mathscr{A} \times \mathscr{B}) \to \overline{\mathbb{R}}_{\geq 0}$ defined by
$$\mu \times \nu(S) = \begin{cases} 0, & S = \emptyset, \\ \infty, & S \neq \emptyset, \end{cases}$$
is a measure and satisfies $\mu \times \nu(A \times B) = \mu(A)\nu(B)$. Note, however, that since μ and ν are not σ-finite, we cannot use Theorem 5.3.33 to assert the existence of this measure except in the trivial case when $X = Y = \emptyset$.
2. Let X and Y be sets with $\mathscr{A}$ and $\mathscr{B}$ σ-algebras on X and Y, respectively. Define $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ and $\nu\colon \mathscr{B} \to \overline{\mathbb{R}}_{\geq 0}$ by
$$\mu(A) = \begin{cases} \operatorname{card}(A), & \operatorname{card}(A) < \infty, \\ \infty, & \text{otherwise}, \end{cases} \qquad \nu(B) = \begin{cases} \operatorname{card}(B), & \operatorname{card}(B) < \infty, \\ \infty, & \text{otherwise}. \end{cases}$$
Then the map $\mu \times \nu\colon \sigma(\mathscr{A} \times \mathscr{B}) \to \overline{\mathbb{R}}_{\geq 0}$ defined by
$$\mu \times \nu(S) = \begin{cases} \operatorname{card}(S), & \operatorname{card}(S) < \infty, \\ \infty, & \text{otherwise}, \end{cases}$$
satisfies $\mu \times \nu(A \times B) = \mu(A)\nu(B)$. By Theorem 5.3.33 we can infer that $\mu \times \nu$ is a measure and is the unique measure with this property. •

5.3.36 Remark (Completeness of product measures) The product measure of complete measure spaces may be incomplete. We shall see a concrete instance of this in Section 5.5.4, but it is revealing to see how this can arise in a general way. Suppose that we have complete measure spaces $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$. Let $A \subseteq X$ be a nonempty set such that $\mu(A) = 0$ (thus A is measurable since $(X, \mathscr{A}, \mu)$ is complete) and let $B \subseteq Y$ be a nonmeasurable set. (Note that it might happen that there are no sets A and B with these properties.) Note that $A \times B \subseteq A \times Y$ and that $A \times Y$ is measurable, being a product of measurable rectangles. Moreover, $\mu \times \nu(A \times Y) = \mu(A)\nu(Y) = 0$ and so $A \times B$ is a subset of a set of measure zero. However, we claim that $A \times B$ is not $\sigma(\mathscr{A} \times \mathscr{B})$-measurable. Indeed, by Proposition 5.2.18, were $A \times B$ to be $\sigma(\mathscr{A} \times \mathscr{B})$-measurable, it would follow that B is $\mathscr{B}$-measurable, which we suppose not to be the case. •

5.3.7 Signed measures

In this section until now, a measure has been thought of as measuring the “size” of a measurable set, and so is an intrinsically nonnegative quantity. However, sometimes one wishes to use measures in ways more subtle than simply to measure “size,” and in this case one wishes to allow for the measure of a set to be negative. In this section we carry out the steps needed to make such a definition, and we give a few basic properties of the sorts of measures we produce. The most interesting examples arise through integration; see Proposition 5.7.65. However, in Theorem 5.3.42 we will characterise signed measures to the degree that it is easy to see exactly what they “are.”

We can begin with the definition.

5.3.37 Definition (Signed measure) For a measurable space $(X, \mathscr{A})$, a ***signed measure*** on $\mathscr{A}$ is a map $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}$ such that

(i) $\mu(\emptyset) = 0$ and

(ii) μ is countably-additive.

A ***signed measure space*** is a triple $(X, \mathscr{A}, \mu)$ where $(X, \mathscr{A})$ is a measurable space and μ is a signed measure on $\mathscr{A}$. •

Note that, by Proposition 5.3.2(viii), a signed measure is consistent, and so a signed measure cannot take both values ∞ and $-\infty$. If, for emphasis, we wish to differentiate between a signed measure and a measure in the sense of Definition 5.3.7, we shall sometimes call the latter a ***positive measure***. However, whenever we say “measure,” we always mean a measure in the sense of Definition 5.3.7.

Let us provide some simple examples of signed measures.

5.3.38 Examples (Signed measures)

1. Let X be a set and let $x_1, x_2 \in X$ be distinct points. Let us take $\mathscr{A} = \mathbf{2}^X$ and define $\mu\colon \mathbf{2}^X \to \overline{\mathbb{R}}$ by
$$\mu(A) = \begin{cases} m_1, & x_1 \in A,\ x_2 \notin A, \\ -m_2, & x_2 \in A,\ x_1 \notin A, \\ m_1 - m_2, & x_1, x_2 \in A, \\ 0, & x_1, x_2 \notin A, \end{cases}$$
for $m_1, m_2 \in \mathbb{R}$. Intuitively, μ has a positive mass m_1 at x_1 and a negative mass $-m_2$ and x_2.

2. Let $X = \mathbb{Z}$ be a set and take $\mathscr{A} = \mathbf{2}^X$. Suppose that the sequences $(p_j)_{j\in\mathbb{Z}_{\geq 0}}$ and $(n_j)_{j\in\mathbb{Z}_{>0}}$ of positive numbers are such that
$$\sum_{j=0}^{\infty} p_j < \infty, \qquad \sum_{j=1}^{\infty} n_j < \infty.$$
For $A \subseteq \mathbb{Z}$ define
$$\mu(A) = \sum_{j\in A\cap\mathbb{Z}_{\geq 0}} p_j - \sum_{j\in A\cap\mathbb{Z}_{<0}} n_{-j},$$

which can easily be verified to define a signed measure. •

Let us now indicate some of the essential features of signed measures.

5.3.39 Definition (Positive and negative sets, Hahn decomposition) For a signed measure space $(X, \mathscr{A}, \mu)$, a set $A \in \mathscr{A}$ is ***positive*** (resp. ***negative***) if, for every $B \subseteq A$ such that $B \in \mathscr{A}$, it holds that $\mu(B) \in \mathbb{R}_{\geq 0}$ (resp. $\mu(B) \in \mathbb{R}_{\leq 0}$). A ***Hahn decomposition*** for $(X, \mathscr{A}, \mu)$ is a pair (P, N) with the following properties:

(i) $P, N \in \mathscr{A}$;

(ii) $X = P \cup N$ and $P \cap N = \emptyset$;

(iii) P is a positive set and N is a negative set. •

It is clear that if A is a positive (resp. negative) set, every measurable subset of A is also positive (resp. negative).

We can prove that Hahn decompositions exist.

5.3.40 Theorem (Hahn Decomposition Theorem) *Every signed measure space possesses a Hahn decomposition. Moreover, if* $(\mathrm{P}_1, \mathrm{N}_1)$ *and* $(\mathrm{P}_2, \mathrm{N}_2)$ *are Hahn decompositions for a signed measure space* $(\mathrm{X}, \mathscr{A}, \mu)$*, then* $\mathrm{P}_1 \cap \mathrm{N}_2$ *and* $\mathrm{P}_2 \cap \mathrm{N}_1$ *both have measure zero.*

Proof Since μ is consistent, we assume without loss of generality that μ cannot take the value $-\infty$. Let us define

$$L = \inf\{\mu(A) \mid A \text{ is a negative set}\}.$$

Note that there are negative sets since $\emptyset$ is negative. Also, $L > -\infty$. Indeed, if $L = -\infty$ this would imply that for each $j \in \mathbb{Z}_{>0}$ there exists a negative set A_j for which $\mu(A_j) < -j$. Let $B_k = \cup_{j=1}^{k} A_k$ so that $B_k \subseteq B_{k+1}$. Note that $\mu(B_k) < -k$. Countable-additivity of μ and Proposition 5.3.3 imply that

$$\mu\Big(\bigcup_{k \in \mathbb{Z}_{>0}} B_k\Big) = \lim_{k \to \infty} \mu(B_k) = -\infty,$$

and so indeed we must have $L > -\infty$ if μ cannot take the value $-\infty$. Now let $(A_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence of sets from $\mathscr{A}$ for which $\lim_{j \to \infty} \mu(A_j) = L$ and define $N = \cup_{j \in \mathbb{Z}_{>0}} A_j$. We claim that N is a negative set. Certainly $N \in \mathscr{A}$, N being a countable union of sets from $\mathscr{A}$. By Lemma 1 from the proof of Proposition 5.3.2 we can write $N = \cup_{j \in \infty} N_j$ for a pairwise disjoint family of negative sets $(N_j)_{j \in \mathbb{Z}_{>0}}$. Now, if $A \subseteq N$ is $\mathscr{A}$-measurable then $A = \cup_{j \in \mathbb{Z}_{>0}} A \cap N_j$. Since $A \cap N_j \subseteq N_j$ it follows that $\mu(A \cap N_j) \in \mathbb{R}_{\leq 0}$. Thus, by countable-additivity of μ,

$$\mu(A) = \sum_{j=1}^{\infty} \mu(A \cap N_j) \leq 0,$$

so showing that N is a negative set.

Now define $P = X \setminus N$. To prove that P is a positive set, we need a lemma.

1 Lemma *If* $(\mathrm{X}, \mathscr{A}, \mu)$ *is a signed measure space and if* $\mathrm{A} \in \mathscr{A}$ *satisfies* $\mu(\mathrm{A}) \in \mathbb{R}_{<0}$*, then there exists a negative set* $\mathrm{B} \subseteq \mathrm{A}$ *such that* $\mu(\mathrm{B}) \leq \mu(\mathrm{A})$*.*

Proof We define a sequence $(m_j)_{j\in\mathbb{Z}_{>0}}$ of nonnegative real numbers and a sequence $(A_j)_{j\in\mathbb{Z}_{>0}}$ of pairwise disjoint $\mathscr{A}$-measurable subsets of A with nonnegative measure as follows. Let

$$m_1 = \sup\{\mu(B) \mid B \in \mathscr{A},\ B \subseteq A\}.$$

Note that $m_1 \in \mathbb{R}_{\geq 0}$ since $\emptyset \in \mathscr{A}$ and $\emptyset \subseteq A$. Now let $A_1 \in \mathscr{A}$ be a subset of A that satisfies $\mu(A_1) \geq \min\{\frac{m_1}{2}, 1\}$, this being possible by the definition of m_1. Note that $\mu(A_1) \in \mathbb{R}_{\geq 0}$. Now suppose that we have defined $m_1, \dots, m_k \in \mathbb{R}_{\geq 0}$ and pairwise disjoint $\mathscr{A}$-measurable sets $A_1, \dots, A_k \subseteq A$ such that $\mu(A_j) \in \mathbb{R}_{\geq 0}$, $j \in \{1, \dots, k\}$. Then let

$$m_{k+1} = \sup\{\mu(B) \mid B \in \mathscr{A},\ B \subseteq A \setminus \cup_{j=1}^{k} A_j\}$$

and let $A_{k+1} \subseteq A \setminus \cup_{j=1}^{k} A_j$ have the property that $\mu(A_{k+1}) \geq \min\{\frac{m_{k+1}}{2}, 1\}$. As we argued above for m_1 and A_1, $m_{k+1}, \mu(A_{k+1}) \in \mathbb{R}_{\geq 0}$. It is clear that $A_{k+1} \cap A_j = \emptyset$, $j \in \{1, \dots, k\}$. Thus $(A_1, \dots, A_{k+1})$ are pairwise disjoint.

Let us take $B = A \setminus \cup_{j\in\mathbb{Z}_{>0}} A_j$. Note that

$$\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j)$$

since the sets $(A_j)_{j\in\mathbb{Z}_{>0}}$ are pairwise disjoint. Therefore,

$$\mu(A) = \mu(B) + \mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) \geq \mu(B).$$

Now we show that B is a negative set. Note that

$$\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j) < \infty.$$

since $|\mu(A)| < \infty$. Thus the sum in the middle converges, and by Proposition 2.4.7 it follows that $\lim_{j\to\infty} \mu(A_j) = 0$. Therefore, $\lim_{j\to\infty} m_j = 0$. Now let $E \subseteq B$ be $\mathscr{A}$-measurable. Thus $E \subseteq A \setminus \cup_{j=1}^{k} A_j$ for every $k \in \mathbb{Z}_{>0}$. Therefore, by definition of m_{k_1}, $\mu(E) \leq m_{k+1}$ for every $k \in \mathbb{Z}_{\geq 0}$. Therefore, it must be the case that $\mu(E) \in \mathbb{R}_{\leq 0}$ since $\lim_{j\to\infty} m_j = 0$. ▼

Now suppose that there exists a set $A \subseteq P$ such that $\mu(A) \in \mathbb{R}_{<0}$. Then, by the lemma, there exists a negative set $B \subseteq A$ such that $\mu(B) \leq \mu(A)$. Now $N \cup B$ is a negative set such that

$$\mu(N \cup B) = \mu(N) + \mu(B) \leq \mu(N) + \mu(A) < \mu(N) = L,$$

which contradicts the definition of L. Thus P is indeed positive.

To prove the final assertion of the theorem, note that both $P_1 \cap N_2$ and $P_2 \cap N_1$ are both positive and negative sets. It must, therefore, be the case that both have measure zero. ■

The Hahn decomposition can be illustrated for our examples above of signed measures.

5.3.41 Examples (Hahn decomposition)

1. We consider Example 5.3.38–1. A Hahn decomposition in this case consists of any subsets P and N such that
   (a) $P \cap N = \emptyset$,
   (b) $P \cup N = X$, and
   (c) $x_1 \in P$ and $x_2 \in N$.

   Note that there will generally be many possible Hahn decompositions in this case, since there are possible many sets of measure zero.
2. For Example 5.3.38–2, a Hahn decomposition is given by $P = \mathbb{Z}_{\geq 0}$ and $N = \mathbb{Z}_{<0}$. If none of the numbers p_j, $j \in \mathbb{Z}_{\geq 0}$, and n_j, $j \in \mathbb{Z}_{>0}$, are zero (as was assumed), then this is the *only* Hahn decomposition. •

As a direct consequence of the Hahn Decomposition Theorem we have the following decomposition of μ.

5.3.42 Theorem (Jordan Decomposition Theorem) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *the following statement hold:*

(i) if ν_+ and ν_- are two positive measures on $\mathscr{A}$, at least one of which is finite, then the map $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}$ defined by $\nu(\mathsf{A}) = \nu_+(\mathsf{A}) - \nu_-(\mathsf{A})$ is a signed measure on $\mathscr{A}$;

(ii) if μ is a signed measure on $\mathscr{A}$ then there exist unique positive measures μ_+ and μ_- on $\mathscr{A}$ such that

(a) at least one of μ_+ and μ_- is finite,

(b) $\mu(\mathsf{A}) = \mu_+(\mathsf{A}) - \mu_-(\mathsf{A})$ for every $\mathsf{A} \in \mathscr{A}$, and

(c) $\mu_+(\mathsf{A}) = \mu(\mathsf{A})$ for every positive set A and $\mu_-(\mathsf{B}) = -\mu(\mathsf{B})$ for every negative set B;

(iii) if ν_+ and ν_- are positive measures on $\mathscr{A}$, at least one of which is finite, such that $\mu(\mathsf{A}) = \nu_+(\mathsf{A}) - \nu_-(\mathsf{A})$ for every $\mathsf{A} \in \mathscr{A}$ and if μ_+ and μ_- are as in part (ii), then $\nu_+(\mathsf{A}) \geq \mu_+(\mathsf{A})$ and $\nu_-(\mathsf{A}) \geq \mu_-(\mathsf{A})$ for every $\mathsf{A} \in \mathscr{A}$.

Proof (i) This is a straightforward verification that ν as defined in the statement of the theorem is countably-additive and satisfies $\mu(\emptyset) = 0$.

(ii) Let (P, N) be a Hahn decomposition for $(X, \mathscr{A}, \mu)$. Note that at most one of the relations $\mu(P) = \infty$ and $\mu(N) = -\infty$ can hold by consistency of μ. Define $\mu_+, \mu_-\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\mu_+(A) = \mu(P \cap A), \quad \mu_-(A) = -\mu(N \cap A).$$

Clearly $\mu_+(\emptyset) = \mu_-(\emptyset) = 0$. Also, for a pairwise disjoint family $(A_j)_{j \in \mathbb{Z}_{>0}}$ of $\mathscr{A}$-measurable sets, we have

$$\mu_+\Big(\bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) = \mu\Big(P \cap \bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) = \mu\Big(\bigcup_{j \in \mathbb{Z}_{>0}} P \cap A_j\Big) = \sum_{j=1}^{\infty} \mu(P \cap A_j) = \sum_{j=1}^{\infty} \mu_+(A_j),$$

giving countable-additivity of μ_+. One similarly shows countable-additivity of μ_-. Also, if $A \in \mathscr{A}$, we have

$$\mu(A) = \mu(P \cap A) + \mu(N \cap A) = \mu_+(A) - \mu_-(A).$$

This gives the existence assertion of this part of the theorem.

We make an observation before we begin the proof of the uniqueness assertion of this part of the theorem. We continue with the notation from the proof of existence above, with μ_+ and μ_- as defined in that part of the proof, relative to the Hahn decomposition (P, N) for $(X, \mathscr{A}, \mu)$. Let (P', N') be another Hahn decomposition. We can then write

$$P = (P' \cap P) \cup (N' \cap P),$$

where, by Theorem 5.3.40, $\mu(N' \cap P) = 0$. Now note that, for every $A \in \mathscr{A}$,

$$\begin{aligned}\mu(P \cap A) &= \mu(((P' \cap P) \cup (N' \cap P)) \cap A) \\ &= \mu(((P' \cap P) \cap A) \cup ((N' \cap P) \cap A)) = \mu(((P' \cap P) \cap A)).\end{aligned}$$

Now we have

$$P' = (P' \cap P) \cup (P' \cap N),$$

where $\mu(P' \cap N) = 0$ by Theorem 5.3.40. Therefore,

$$\mu(P' \cap A) = \mu((P' \cap P) \cap A),$$

from which we deduce that $\mu(P \cap A) = \mu(P' \cap A)$. Similarly, we show that $\mu(N \cap A) = \mu(N' \cap A)$.

Let μ'_+ and μ'_- be positive measures satisfying

$$\mu(A) = \mu'_+(A) - \mu'_-(A), \qquad A \in \mathscr{A},$$

and suppose that $\mu'_+(A) = \mu(A)$ for every positive set A and that $\mu'_-(B) = \mu(B)$ for every negative set B. Let $A \in \mathscr{A}$ be a positive set and let $B \in \mathscr{A}$ be a negative set. Then, for the Hahn decomposition (P, N), we write

$$A = (P \cap A) \cup (N \cap A).$$

Since A is a positive set, we must have $\mu(N \cap A) = 0$. Define $P' = P \cup (N \cap A)$ and $N' = X \setminus P'$. Obviously P' is a positive set, being the union of a positive set with a set of measure zero. Since $N' = N \setminus (N \cap A)$, it follows that N' is a negative set. Thus (P', N') is a Hahn decomposition. Moreover, $P' \cap A = A$, and so

$$\mu'_+(A) = \mu(P' \cap A) = \mu(P \cap A) = \mu_+(A),$$

the second equality following from the remarks beginning this part of the proof. Similarly one shows that $\mu'_-(B) = \mu_-(B)$. Thus any positive measures μ'_+ and μ'_- having the three stated properties must agree with the measures μ_+ and μ_- explicitly constructed in part (ii).

(iii) For a positive set A we have

$$\mu(A) = \mu_+(A) = \nu_+(A) - \nu_-(A)$$

and so $\nu_+(A) \geq \mu_+(A)$ for every positive set A. For a negative set B we have $\mu_+(B) = 0$ and so we immediately have $\nu_+(B) \geq \mu_+(B)$. Therefore, for $A \in \mathscr{A}$ we have

$$\begin{aligned}& A = (P \cap A) \cup (N \cap A) \\ \implies \quad & \nu_+(A) = \nu_+(P \cap A) + \nu_+(N \cap A) \geq \mu_+(P \cap A) + \mu_+(N \cap A) = \mu_+(A).\end{aligned}$$

By the same arguments, *mutatis mutandis*, one shows that $\nu_-(A) \geq \mu_-(A)$ for every $A \in \mathscr{A}$. ■

Note that, without all of the assumptions from part (ii) of the theorem, uniqueness of μ_+ and μ_- cannot be guaranteed. Indeed, if μ is a positive measure then we can write

$$\mu(A) = \mu_+(A) - \mu_-(A) = \nu_+(A) - \nu_-(B)$$

where $\mu_+ = \mu$, μ_- is the zero measure, $\nu_+ = 2\mu$, and $\nu_- = \mu$. Note that $\nu_+(A) \geq \mu_+(A)$ and $\nu_-(A) \geq \mu_-(A)$, as asserted in part (iii).

Thus we make the following definition.

5.3.43 Definition (Jordan decomposition) If $(X, \mathscr{A}, \mu)$ is a signed measure space, the ***positive part*** and the ***negative part*** of μ are the positive measures μ_+ and μ_-, respectively, having the following properties:

(i) $\mu(A) = \mu_+(A) - \mu_-(A)$ for every $A \in \mathscr{A}$;

(ii) $\mu'_+(A) = \mu(A)$ for every positive set A;

(iii) $\mu'_-(B) = -\mu(B)$ for every negative set B.

The ***Jordan decomposition*** of μ is given by the representation $\mu = \mu_+ - \mu_-$ which signifies the first of the above properties of μ_+ and μ_-. •

5.3.44 Remark (Connections to functions with bounded variation) In Theorem 3.3.3 we considered the Jordan decomposition for a function of bounded variation. This decomposition, like the one in Theorem 5.3.42, gives an additive decomposition with a (sort of) positive component and a (sort of) negative component. There is, as one might hope, a concrete relationship between the two Jordan decompositions. However, this will not be realised until ***missing stuff***. •

For our ongoing examples we can illustrate the Jordan decomposition.

5.3.45 Examples (Jordan decomposition)

1. For the signed measure of Example 5.3.38–1, the positive and negative parts of the signed measure μ are defined by

$$\mu_+(A) = \begin{cases} m_1, & x_1 \in A, \\ 0, & x_1 \notin A, \end{cases} \quad \mu_-(A) = \begin{cases} m_2, & x_2 \in A, \\ 0, & x_2 \notin A. \end{cases}$$

2. For the signed measure of Example 5.3.38–2, the positive and negative parts of the signed measure μ are defined by

$$\mu_+(A) = \begin{cases} \sum_{j \in A \cap \mathbb{Z}_{\geq 0}} p_j, & A \cap \mathbb{Z}_{\geq 0} \neq \emptyset, \\ 0, & A \cap \mathbb{Z}_{\geq 0} = \emptyset, \end{cases} \quad \mu_-(A) = \begin{cases} \sum_{j \in A \cap \mathbb{Z}_{<0}} n_{-j}, & A \cap \mathbb{Z}_{<0} \neq \emptyset, \\ 0, & A \cap \mathbb{Z}_{<0} = \emptyset. \end{cases}$$

•

Now that we have at hand the decompositions which we use to characterise signed measures, we can use these to provide a new measure associated with a signed measure. The value of this construction may not be immediately apparent, but will be made clear in ***missing stuff***.

5.3.46 Definition (Variation and total variation of a signed measure) For a signed measure space $(X, \mathscr{A}, \mu)$, the ***variation*** of μ is the positive measure $|\mu|\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$|\mu|(A) = \mu_+(A) + \mu_-(A),$$

where μ_+ and μ_- are the positive and negative parts, respectively, μ. The ***total variation*** of μ is $\|\mu\| = |\mu|(X)$ •

It is a simple verification to check that $|\mu|$ is indeed a positive measure. The following result characterises it among all positive measures which relate to μ in a prescribed manner.

5.3.47 Proposition (Property of the variation of a signed measure) *For* $(X, \mathscr{A}, \mu)$ *a signed measure space,* $|\mu(A)| \leq |\mu|(A)$ *for all* $A \in \mathscr{A}$*. Moreover, if* $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ *is a positive measure such that* $|\mu(A)| \leq \nu(A)$ *for every* $A \in \mathscr{A}$*, then* $|\mu|(A) \leq \nu(A)$ *for every* $A \in \mathscr{A}$*.*

Proof The first assertion of the result is clear, for if $A \in \mathscr{A}$ then

$$|\mu(A)| = |\mu_+(A) - \mu_-(A)| \leq \mu_+(A) + \mu_-(A) = |\mu|(A).$$

For the second assertion, suppose that ν is a positive measure with the property that $|\mu(A)| \leq \nu(A)$ for every $A \in \mathscr{A}$. If (P, N) is a Hahn decomposition for $(X, \mathscr{A}, \mu)$ then, for any $A \in \mathscr{A}$,

$$\mu_+(P \cap A) = |\mu(P \cap A)| \leq \nu(P \cap A)$$

and

$$\mu_-(N \cap A) = |\mu(N \cap A)| \leq \nu(N \cap A).$$

Therefore, using the definition of μ_+ and μ_-,

$$|\mu|(A) = \mu_+(A) + \mu_-(A) = \mu_+(P \cap A) + \mu_-(N \cap A) \leq \nu(P \cap A) + \nu(N \cap A) = \nu(A),$$

as desired. ■

The following property of the variation of a signed measure is also useful.

5.3.48 Proposition (Characterisation of the variation of a signed measure) *For a signed measure space* $(X, \mathscr{A}, \mu)$ *and for* $A \in \mathscr{A}$*,*

$$|\mu|(A) = \sup\Big\{\sum_{j=1}^{k} |\mu(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } A\Big\}.$$

Proof Let $A \in \mathscr{A}$. For a partition $(A_1, \dots, A_k)$ of A we have

$$|\mu|(A) = \sum_{j=1}^{k} |\mu|(A_j) \geq \sum_{j=1}^{k} |\mu(A_j)|.$$

by Proposition 5.3.47 and using countable-additivity (and hence finite-additivity) of $|\mu|$. Taking the supremum of the expression on the right over all partitions gives

$$|\mu|(A) \geq \sup\Big\{\sum_{j=1}^{k} |\mu(B_j)| \;\Big|\; (B_1, \dots, B_k) \text{ is a partition of } A\Big\}.$$

We also have, for a Hahn decomposition (P, N) for $(X, \mathscr{A}, \mu)$ and a partition $(A_1, \dots, A_k)$ for A,

$$\mu_+(P \cap A) = |\mu(P \cap A)| = \Big|\sum_{j=1}^{k} \mu(P \cap A_j)\Big| \leq \sum_{j=1}^{k} |\mu(P \cap A_j)|$$

and similarly

$$\mu_-(N \cap A) \leq \sum_{j=1}^{k} |\mu(N \cap A_j)|.$$

Therefore, using the definition of μ_+ and μ_-,

$$|\mu|(A) = \mu_+(P \cap A) + \mu_-(N \cap A) \leq \sum_{j=1}^{k} |\mu(P \cap A_j)| + \sum_{j=1}^{k} |\mu(N \cap A_j)|.$$

Since $(P \cap A_1, \dots, P \cap A_k, N \cap A_1, \dots, N \cap A_k)$ is a partition of A we have

$$|\mu|(A) \leq \sup\Big\{\sum_{j=1}^{k} |\mu(B_j)| \;\Big|\; (B_1, \dots, B_k) \text{ is a partition of } A\Big\},$$

which gives the result. ■

The total variation is, in fact, an interesting quantity; it is a norm on the set of finite signed measures. This point of view will be taken up in Section 6.7.9.

As with measures, we can restrict signed measures to measurable subsets.

5.3.49 Proposition (Restriction of a signed measure) *If* $(\mathrm{X}, \mathscr{A}, \mu)$ *is a signed measure space and if* $\mathrm{A} \in \mathscr{A}$*, then* $(\mathrm{A}, \mathscr{A}_\mathrm{A}, \mu|\mathscr{A}_\mathrm{A})$ *is a signed measure space. (See Proposition 5.2.6 for the definition of* $\mathscr{A}_\mathrm{A}$*.)*

Proof This follows very much along the lines of Proposition 5.3.18. ■

5.3.8 Complex measures

Next we consider measures taking not just general real values, but complex values. As with signed measures, we shall not be able to see interesting examples of complex measures until we talk about integration; see Proposition 5.7.65.

We begin with the definition.

5.3.50 Definition (Complex measure) For a measurable space $(X, \mathscr{A})$, a ***signed measure*** on $\mathscr{A}$ is a map $\mu\colon \mathscr{A} \to \mathbb{C}$ such that

(i) $\mu(\emptyset) = 0$ and

(ii) $\mu\Big(\bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j)$ for every family $(A_j)_{j \in \mathbb{Z}_{>0}}$ of pairwise disjoint sets from $\mathscr{A}$ (***countable-additivity***).

A ***complex measure space*** is a triple $(X, \mathscr{A}, \mu)$ where $(X, \mathscr{A})$ is a measurable space and μ is a complex measure on $\mathscr{A}$. •

Note that a complex measure is intrinsically finite since it must take values in $\mathbb{C}$. This makes complex measures a little different and more restrictive in scope than positive or signed measures.

For a complex measure space $(X, \mathscr{A}, \mu)$, we can define finite signed measures $\mathrm{Re}(\mu), \mathrm{Im}(\mu)\colon \mathscr{A} \to \mathbb{R}$ by

$$\mathrm{Re}(\mu)(A) = \mathrm{Re}(\mu(A)), \quad \mathrm{Im}(\mu)(A) = \mathrm{Im}(\mu(A)), \qquad A \in \mathscr{A}.$$

We obviously call $\mathrm{Re}(\mu)$ the ***real part*** of μ and $\mathrm{Im}(\mu)$ the ***imaginary part*** of μ. It is trivial to verify that $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$ are indeed finite signed measures, and the reader can do this as Exercise 5.3.5. We can then write

$$\mu(A) = \mathrm{Re}(\mu)(A) + \mathrm{i}\,\mathrm{Im}(\mu)(A),$$

or $\mu = \mathrm{Re}(\mu) + \mathrm{i}\,\mathrm{Im}(\mu)$ for short. Since $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$ are signed measures, they have Jordan decompositions

$$\mathrm{Re}(\mu) = \mathrm{Re}(\mu)_+ - \mathrm{Re}(\mu)_-, \quad \mathrm{Im}(\mu) = \mathrm{Im}(\mu)_+ - \mathrm{Im}(\mu)_-.$$

We can then write

$$\mu = \mathrm{Re}(\mu)_+ - \mathrm{Re}(\mu)_- + \mathrm{i}(\mathrm{Im}(\mu)_+ - \mathrm{Im}(\mu)_-),$$

to which we refer as the ***Jordan decomposition*** of the complex measure μ. It is clear that a finite signed measure can be thought of as a complex measure whose imaginary part is the zero measure.

Now let us turn to the variation of a complex measure.

5.3.51 Definition (Variation and total variation of a complex measure) Let $(X, \mathscr{A}, \mu)$ be a complex measure space. The ***variation*** of μ is the map $|\mu|\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$|\mu|(A) = \sup\Big\{\sum_{j=1}^{k} |\mu(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } A\Big\}.$$

The ***total variation*** of μ is $\|\mu\| = |\mu|(X)$. •

Different from the case of a signed measure, it is not immediately clear that the variation is a measure. Thus we verify this.

5.3.52 Proposition (Variation is a positive finite measure) *If $(\mathrm{X}, \mathscr{A}, \mu)$ is a complex measure space then $|\mu|$ is a finite positive measure that satisfies $|\mu(\mathrm{A})| \leq |\mu|(\mathrm{A})$ for every $\mathrm{A} \in \mathscr{A}$. Moreover, if $\nu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ is a positive measure satisfying $|\mu(\mathrm{A})| \leq \nu(\mathrm{A})$ for every $\mathrm{A} \in \mathscr{A}$, then $|\mu|(\mathrm{A}) \leq \nu(\mathrm{A})$ for every $\mathrm{A} \in \mathscr{A}$.*

Proof It is evident that $|\mu|(\emptyset) = 0$. To verify countable-additivity of $|\mu|$, we first verify finite-additivity. Let $A_1, A_2 \in \mathscr{A}$ be disjoint and let $(B_1, \dots, B_k)$ be a partition of $A_1 \cup A_2$.

We then have

$$\begin{aligned}\sum_{j=1}^{k}|\mu(B_j)| &= \sum_{j=1}^{k}|\mu(A_1\cap B_j)+\mu(A_2\cap B_j)|\\ &\le \sum_{j=1}^{k}(|\mu(A_1\cap B_j)|+|\mu(A_2\cap B_j)|)\le|\mu|(A_1)+|\mu|(A_2),\end{aligned}$$

the last inequality by definition of $|\mu|$. Since

$$|\mu|(A_1\cup A_2)=\sup\Big\{\sum_{j=1}^{k}|\mu(B_j)|\;\Big|\;(B_1,\dots,B_k)\text{ is a partition of }A_1\cup A_2\Big\},$$

we have

$$|\mu|(A_1\cup A_2)\le|\mu|(A_1)+|\mu|(A_2).$$

Now let $(B_{1,1},\dots,B_{1,k_1})$ be a partition of A_1 and let $(B_{2,1},\dots,B_{2,k_2})$ be a partition of A_2. Since

$$(B_{1,1},\dots,B_{1,k_1})\cup(B_{2,1},\dots,B_{2,k_2})$$

is a partition of $A_1\cup A_2$ we have

$$\sum_{j_1=1}^{k_1}|\mu(B_{1,j_1})|+\sum_{j_2=1}^{k_2}|\mu(B_{2,j_2})|\le|\mu|(A_1\cup A_2).$$

Since

$$\begin{aligned}|\mu|(A_1)&=\sup\Big\{\sum_{j_1=1}^{k_1}|\mu(B_{1,j_1})|\;\Big|\;(B_{1,1},\dots,B_{1,k_1})\text{ is a partition of }A_1\Big\},\\ |\mu|(A_2)&=\sup\Big\{\sum_{j_2=1}^{k_2}|\mu(B_{2,j_2})|\;\Big|\;(B_{2,1},\dots,B_{2,k_2})\text{ is a partition of }A_2\Big\}\end{aligned}$$

we have

$$|\mu|(A_1)+|\mu|(A_2)\le|\mu|(A_1\cup A_2).$$

Thus $|\mu|(A_1\cup A_2)=|\mu|(A_1)+|\mu|(A_2)$, whence follows the finite additivity of $|\mu|$.

Now note that for $A\in\mathscr{A}$ we have

$$|\mu(A)|\le|\mathrm{Re}(\mu)(A)|+|\mathrm{Im}(\mu)(A)| \tag{5.6}$$

by ***missing stuff***. Therefore, for $A\in\mathscr{A}$ and for a finite partition $(A_1,\dots,A_k)$ for A, we have

$$\begin{aligned}\sum_{j=1}^{k}|\mu(A_j)| &\le \sum_{j=1}^{k}(|\mathrm{Re}(\mu)(A_j)|+|\mathrm{Im}(\mu)(A_j)|)\\ &\le \sum_{j=1}^{k}(\mathrm{Re}(\mu)_+(A_j)+\mathrm{Re}(\mu)_-(A_j)+\mathrm{Im}(\mu)_+(A_j)+\mathrm{Im}(\mu)_-(A_j))\\ &= \sum_{j=1}^{k}\mathrm{Re}(\mu)_+(A)+\mathrm{Re}(\mu)_-(A)+\mathrm{Im}(\mu)_+(A)+\mathrm{Im}(\mu)_-(A).\end{aligned}$$

Taking the supremum of the leftmost expression over all partitions we have

$$|\mu|(A) \leq \operatorname{Re}(\mu)_+(A) + \operatorname{Re}(\mu)_-(A) + \operatorname{Im}(\mu)_+(A) + \operatorname{Im}(\mu)_-(A). \tag{5.7}$$

Therefore, if $(A_j)_{j\in\mathbb{Z}_{>0}}$ is a sequence of sets from $\mathscr{A}$ having the properties that $A_j \supseteq A_{j+1}$, $j \in \mathbb{Z}_{>0}$, and that $\cap_{j\in\mathbb{Z}_{>0}} A_j = \emptyset$, we have

$$\lim_{j\to\infty} |\mu|(A_j) \leq \lim_{j\to\infty}(\operatorname{Re}(\mu)_+(A_j) + \operatorname{Re}(\mu)_-(A_j) + \operatorname{Im}(\mu)_+(A_j) + \operatorname{Im}(\mu)_-(A_j)) = 0.$$

Countable-additivity of $|\mu|$ now follows from Proposition 5.3.3.

The finiteness of $|\mu|$ follows immediately from (5.7), noting that the four positive measures on the right are finite.

For $A \in \mathscr{A}$ and for a partition $(A_1, \dots, A_k)$ of A we have

$$|\mu(A)| \leq \sum_{j=1}^{k} |\mu(A_j)| \leq |\mu|(A),$$

which gives the stated property of $|\mu|$.

Now suppose that ν is a positive measure on $\mathscr{A}$ for which $|\mu(A)| \leq \nu(A)$ for every $A \in \mathscr{A}$. Therefore, for $A \in \mathscr{A}$ and for a partition $(A_1, \dots, A_k)$ of A, we have

$$\sum_{j=1}^{k} |\mu(A_j)| \leq \sum_{j=1}^{k} \nu(A_j) = \nu(A).$$

Taking the supremum of the left-hand side over all partitions then gives $|\mu|(A) \leq \nu(A)$, as desired. ■

Note that Proposition 5.3.48 ensures that if a finite signed measure μ is regarded as a complex measure with zero imaginary part, the definition of $|\mu|$ agrees when defined thinking of μ as a signed measure and when defined thinking of μ as a complex measure.

As with signed measures, the total variation for a complex measure is interesting, and will be studied in Section 6.7.9.

5.3.9 Vector measures

The development of vector measures follows rather like that for complex measures in the preceding section. While it is possible to consider measures taking values in general vector spaces, in this section we restrict ourselves to $\mathbb{R}^n$-valued measures.

5.3.53 Definition (Vector measure) For a measurable space $(X, \mathscr{A})$, a ***vector measure*** on $\mathscr{A}$ is a map $\boldsymbol{\mu}\colon \mathscr{A} \to \mathbb{R}^n$ such that

(i) $\boldsymbol{\mu}(\emptyset) = \mathbf{0}$ and

(ii) $\boldsymbol{\mu}\Big(\bigcup_{j\in\mathbb{Z}} A_j\Big) = \sum_{j=1}^{\infty} \boldsymbol{\mu}(A_j)$ for every family $(A_j)_{j\in\mathbb{Z}_{>0}}$ of pairwise disjoint sets from $\mathscr{A}$ (***countable-additivity***).

A ***vector measure space*** is a triple $(X, \mathscr{A}, \boldsymbol{\mu})$ where $(X, \mathscr{A})$ is a measurable space and $\boldsymbol{\mu}$ is a vector measure on $\mathscr{A}$. •

For a vector measure space $(X, \mathscr{A}, \boldsymbol{\mu})$ with $\boldsymbol{\mu}$ taking values in $\mathbb{R}^n$ and for $j \in \{1, \dots, n\}$ we can define a finite signed measure μ_j by $\mu_j(A) = \mathrm{pr}_j(\boldsymbol{\mu}(A))$, where $\mathrm{pr}_j \colon \mathbb{R}^n \to \mathbb{R}$ is the projection onto the jth component. We can write

$$\boldsymbol{\mu}(A) = \mu_j(A)\boldsymbol{e}_1 + \dots + \mu_n(A)\boldsymbol{e}_n, \qquad A \in \mathscr{A},$$

where $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is the standard basis for $\mathbb{R}^n$. Of course, we can also decompose each of the signed measures $\mu_1, \dots, \mu_n$ into its positive and negative parts, and so arrive at the ***Jordan decomposition*** of $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \mu_{1,+} - \mu_{1,-} + \dots + \mu_{n,+} - \mu_{n,-}.$$

The definition of the variation for vector measures mirrors that for complex measures.

5.3.54 Definition (Variation and total variation of a vector measure) Let $(X, \mathscr{A}, \boldsymbol{\mu})$ be a vector measure space with $\boldsymbol{\mu}$ taking values in $\mathbb{R}^n$. The ***variation*** of $\boldsymbol{\mu}$ is the map $\|\boldsymbol{\mu}\|_{\mathbb{R}^n} \colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$\|\boldsymbol{\mu}\|_{\mathbb{R}^n}(A) = \sup\Big\{ \sum_{j=1}^{k} \|\boldsymbol{\mu}(A_j)\|_{\mathbb{R}^n} \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } A \Big\}.$$

The ***total variation*** of $\boldsymbol{\mu}$ is $|||\boldsymbol{\mu}|||_{\mathbb{R}^n} = \|\boldsymbol{\mu}\|_{\mathbb{R}^n}(X)$. •

As with complex measures, one can verify that the variation of a vector measure defines a positive measure.

5.3.55 Proposition (Variation is a positive finite measure) *If* $(\mathrm{X}, \mathscr{A}, \boldsymbol{\mu})$ *is a vector measure space with* $\boldsymbol{\mu}$ *taking values in* $\mathbb{R}^{\mathrm{n}}$*, then* $\|\boldsymbol{\mu}\|_{\mathbb{R}^{\mathrm{n}}}$ *is a finite positive measure that satisfies* $\|\boldsymbol{\mu}(\mathrm{A})\|_{\mathbb{R}^{\mathrm{n}}} \leq \|\boldsymbol{\mu}\|_{\mathbb{R}^{\mathrm{n}}}(\mathrm{A})$ *for every* $\mathrm{A} \in \mathscr{A}$*. Moreover, if* $\nu \colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ *is a positive measure satisfying* $\|\boldsymbol{\mu}(\mathrm{A})\|_{\mathbb{R}^{\mathrm{n}}} \leq \nu(\mathrm{A})$ *for every* $\mathrm{A} \in \mathscr{A}$*, then* $\|\boldsymbol{\mu}\|_{\mathbb{R}^{\mathrm{n}}}(\mathrm{A}) \leq \nu(\mathrm{A})$ *for every* $\mathrm{A} \in \mathscr{A}$*.*

Proof The proof is very similar to the corresponding Proposition 5.3.52 for complex measures, so we skip the details of the computations, only pointing out the important differences with the previous proof.

It is still clear that $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}(\emptyset) = 0$. The proof of finite-additivity of $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}$ follows in exactly the same manner as the complex case, but with the complex modulus $|\cdot|$ being replaced by the Euclidean norm $\|\cdot\|_{\mathbb{R}^n}$. In the proof of countable-additivity, the relation (5.6) in the complex case is replaced with the relation

$$\|\boldsymbol{\mu}(A)\|_{\mathbb{R}^n} \leq \sum_{j=1}^{n} |\mu_j(A)|,$$

following Proposition **??**. This results in the relation (5.7) in the complex case being replaced with the relation

$$\|\boldsymbol{\mu}(A)\|_{\mathbb{R}^n} \leq \sum_{j=1}^{n} (\mu_{j,+}(A) + \mu_{j,-}(A)) = \sum_{j=1}^{n} |\mu_j|(A). \tag{5.8}$$

Then the proof of countable additivity, using Proposition 5.3.3, follows just as in the complex case, as does finiteness of $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}$.

The property for $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}$ in the proposition is proved just as in the complex case: for $A \in \mathscr{A}$ and for a partition $(A_1, \dots, A_k)$ for A, we have

$$\|\boldsymbol{\mu}(A)\|_{\mathbb{R}^n} \leq \sum_{j=1}^{k} \|\boldsymbol{\mu}(A_j)\|_{\mathbb{R}^n} \leq \|\boldsymbol{\mu}\|_{\mathbb{R}^n}(A).$$

If ν is a positive measure such that $\|\boldsymbol{\mu}(A)\|_{\mathbb{R}^n} \leq \nu(A)$ for every $A \in \mathscr{A}$, we have, just as in the complex case, for a partition $(A_1, \dots, A_k)$ of A:

$$\sum_{j=1}^{k} \|\boldsymbol{\mu}(A_j)\|_{\mathbb{R}^n} \leq \sum_{j=1}^{k} \nu(A_j) = \nu(A),$$

and taking the supremum of the left-hand side over all partitions gives $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}(A) \leq \nu(A)$. ■

5.3.10 Spaces of positive, signed, complex, and vector measures

In this section we briefly consider the various spaces of measures on a measurable space $(X, \mathscr{A})$. Further structural properties of these spaces will be explored in Section 6.7.9.

5.3.56 Definition (Spaces of positive, signed, complex, and vector measures) For a measurable space $(X, \mathscr{A})$, we use the following notation:

(i) $\mathsf{M}((X, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ is the set of positive measures on $\mathscr{A}$;

(ii) $\mathsf{M}((X, \mathscr{A}); \overline{\mathbb{R}})$ is the set of signed measures on $\mathscr{A}$;

(iii) $\mathsf{M}((X, \mathscr{A}); \mathbb{R})$ is the set of finite signed measures on $\mathscr{A}$;

(iv) $\mathsf{M}((X, \mathscr{A}); \mathbb{C})$ is the set of complex measures on $\mathscr{A}$;

(v) $\mathsf{M}((X, \mathscr{A}); \mathbb{R}^n)$ is the set of vector measures on $\mathscr{A}$ taking values in $\mathbb{R}^n$.

For brevity, we may use $\mathsf{M}(X; \overline{\mathbb{R}}_{\geq 0}), \dots, \mathsf{M}(X; \mathbb{R}^n)$ if the σ-algebra $\mathscr{A}$ is understood. •

Let us first explore the algebraic structure of these spaces of measures.

5.3.57 Proposition (The vector space structure of spaces of measures) *For a measurable space* $(\mathrm{X}, \mathscr{A})$*, the following statements hold:*

(i) the set $\mathsf{M}((\mathrm{X}, \mathscr{A}); \mathbb{R})$ *has a* $\mathbb{R}$*-vector space structure with vector addition and scalar multiplication, respectively, defined by*

$$(\mu_1 + \mu_2)(\mathrm{A}) = \mu_1(\mathrm{A}) + \mu_2(\mathrm{A}), \quad (\mathrm{a}\mu)(\mathrm{A}) = \mathrm{a}(\mu(\mathrm{A}))$$

for measures μ*,* μ_1*, and* μ_2 *in* $\mathsf{M}((\mathrm{X}, \mathscr{A}); \mathbb{R})$*, and for* $\mathrm{a} \in \mathbb{R}$*;*

(ii) the set $\mathsf{M}((\mathrm{X}, \mathscr{A}); \mathbb{R}^{\mathrm{n}})$ *has a* $\mathbb{R}$*-vector space structure with vector addition and scalar multiplication, respectively, defined by*

$$(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)(\mathrm{A}) = \boldsymbol{\mu}_1(\mathrm{A}) + \boldsymbol{\mu}_2(\mathrm{A}), \quad (\mathrm{a}\boldsymbol{\mu})(\mathrm{A}) = \mathrm{a}(\boldsymbol{\mu}(\mathrm{A}))$$

for measures $\boldsymbol{\mu}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathsf{M}((\mathrm{X}, \mathscr{A}); \mathbb{R}^{\mathrm{n}})$ *and for* $\mathrm{a} \in \mathbb{R}$*;*

(iii) the set $\mathsf{M}((\mathsf{X},\mathscr{A});\mathbb{C})$ *has a* $\mathbb{C}$*-vector space structure with vector addition and scalar multiplication, respectively, defined by*

$$(\mu_1+\mu_2)(\mathsf{A})=\mu_1(\mathsf{A})+\mu_2(\mathsf{A}),\quad (\mathsf{a}\mu)(\mathsf{A})=\mathsf{a}(\mu(\mathsf{A}))$$

for measures $\mu,\mu_1,\mu_2\in\mathsf{M}((\mathsf{X},\mathscr{A});\mathbb{C})$ *and for* $\mathsf{a}\in\mathbb{C}$.

Proof To check that $\mu_1+\mu_2$ and $a\mu$ (or $\boldsymbol{\mu}_1+\boldsymbol{\mu}_2$ and $a\boldsymbol{\mu}$) have the properties of a measure is straightforward. The remainder of the proof is just a matter of verifying the vector space axioms. The reader who believes this verification might be interesting is welcomed to perform it. ■

5.3.58 Remark (Vector space structures for infinite-valued measures) The reader will have noticed the absence from the above list the vector space structures for the set of positive measures and the set of signed measures. This absence is deserved since, using the natural vector space operations from the statement of the proposition, these sets of measures do not have vector space structures. Let us be sure we understand why in each case.

1. $\mathsf{M}((X,\mathscr{A});\overline{\mathbb{R}}_{\geq 0})$ is not a $\mathbb{R}$-vector space. The problem here is not just the fact that we allow infinite values for the measures. Even if we restrict to finite positive measures, we do not have a natural vector space structure for which vector addition is given by
$$(\mu_1+\mu_2)(A)=\mu_1(A)+\mu_2(A).$$
To see this, let μ be a finite positive measure on $\mathscr{A}$. In order for the operation above to be vector space addition, there must exist a finite positive measure $-\mu$ on $\mathscr{A}$ such that $\mu+(-\mu)$ is the zero measure. Thus, for example, we would have to have $\mu(X)+(-\mu(X))=0$ and so $-\mu(X)\in\mathbb{R}_{<0}$ if $\mu(X)\in\mathbb{R}_{>0}$. In particular, $-\mu$ cannot be a positive measure.
2. $\mathsf{M}((X,\mathscr{A});\overline{\mathbb{R}})$ is not a $\mathbb{R}$-vector space. Indeed, if $(X,\mathscr{A},\mu_1)$ and $(X,\mathscr{A},\mu_2)$ are signed measure spaces for which μ_1 takes the value ∞ and μ_2 takes the value $-\infty$, then (cf. the proof of Proposition 5.3.2(viii)) it follows that $\mu_1(X)=\infty$ and $\mu_2(X)=-\infty$. Therefore, $(\mu_1+\mu_2)(X)$ cannot be defined in the natural way. •

Despite the fact that $\mathsf{M}((X,\mathscr{A});\overline{\mathbb{R}}_{\geq 0})$ is not a $\mathbb{R}$-vector space, we would like for it to have some structure since it comprises the set of positive measures on $\mathscr{A}$, and as such is an interesting object. The following result says that this set is, in fact, a convex cone.

5.3.59 Proposition (The set of positive measures is a convex cone) *Let* $(\mathsf{X},\mathscr{A})$ *be a measurable space, let* $\mu,\mu_1,\mu_2\in\mathsf{M}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}}_{\geq 0})$*, and let* $\mathsf{a}\in\mathbb{R}_{\geq 0}$*. Then the maps* $\mathsf{a}\mu,\mu_1+\mu_2\colon\mathscr{A}\to\overline{\mathbb{R}}_{\geq 0}$ *defined by*

$$(\mathsf{a}\mu)(\mathsf{A})=\mathsf{a}(\mu(\mathsf{A})),\quad (\mu_1+\mu_2)(\mathsf{A})=\mu_1(\mathsf{A})+\mu_2(\mathsf{A})$$

are positive measures on $\mathscr{A}$*. Moreover, for every* $\mu,\mu_1,\mu_2,\mu_3\in\mathsf{M}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}}_{\geq 0})$ *and for every* $\mathsf{a},\mathsf{a}_1,\mathsf{a}_2\in\mathbb{R}_{\geq 0}$*, the following statements hold:*

(i) $\mu_1 + \mu_2 = \mu_2 + \mu_1$;

(ii) $\mu_1 + (\mu_2 + \mu_3) = (\mu_1 + \mu_2) + \mu_3$;

(iii) $\mathrm{a}_1(\mathrm{a}_2\mu) = (\mathrm{a}_1\mathrm{a}_2)\mu$;

(iv) $\mathrm{a}(\mu_1 + \mu_2) = \mathrm{a}\mu_1 + \mathrm{a}\mu_2$;

(v) $(\mathrm{a}_1 + \mathrm{a}_2)\mu = \mathrm{a}_1\mu + \mathrm{a}_2\mu$.

Proof As with the proof of Proposition 5.3.57, the verification of the statements are simple matters of checking the properties. ■

5.3.11 Notes

Exercises

5.3.1 Let X be a set and let $\mathscr{A} \subseteq \mathbf{2}^X$ be an algebra. Let $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ have the property that $\mu(\emptyset) = 0$. Show that μ is countably-additive if and only if it is finitely-additive and countably-subadditive.

5.3.2 Let X be a countable set, let $\mathscr{A}$ be the algebra $\mathscr{A} = \mathbf{2}^X$, and define $\mu\colon \mathbf{2}^X \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\mu(A) = \begin{cases} 0, & \operatorname{card}(A) < \infty, \\ \infty, & \operatorname{card}(A) = \infty. \end{cases}$$

Answer the following questions.

(a) Show that μ is a σ-finite, finitely-additive measure.

(b) Show that if $(A_j)_{j\in\mathbb{Z}_{>0}}$ if a sequence of subsets from $\mathscr{A}$ for which $A_j \supseteq A_{j+1}$, $j \in \mathbb{Z}_{>0}$, for which $\cap_{j\in\mathbb{Z}_{>0}} A_j = \emptyset$, and for which $\mu(A_k) < \infty$ for some $k \in \mathbb{Z}_{>0}$, it holds that $\lim_{j\to\infty} \mu(A_j) = 0$.

(c) Show that μ is not countably-additive.

5.3.3 Let X be a set and consider the collection $\mathscr{S}$ of subsets of X defined by $\mathscr{S} = \{\emptyset\}$. Define $\mu_0\colon \mathscr{S} \to \overline{\mathbb{R}}_{\geq 0}$ by $\mu_0(\emptyset) = 0$. Compute the outer measure generated by $(\mathscr{S}, \mu_0)$.

5.3.4 For a measure space $(X, \mathscr{A}, \mu)$ do the following.

(a) Show that if $(A_j)_{j\in\mathbb{Z}_{>0}}$ is a countable collection of sets of measure zero then

$$\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) = 0.$$

(b) When will there exist an uncountable collection of sets of measure zero whose union has positive measure.

5.3.5 For a complex measure space $(X, \mathscr{A}, \mu)$, show that $\operatorname{Re}(\mu)$ and $\operatorname{Im}(\mu)$ are finite signed measures on $\mathscr{A}$.

5.3.6 Let $(X, \mathscr{A}, \boldsymbol{\mu})$ be a vector measure space with $\boldsymbol{\mu}$ taking values in $\mathbb{R}^n$. Show that for $A \in \mathscr{A}$ we have

$$\sum_{l=1}^{n} |\mu_l|(A) \leq \sqrt{n}\|\boldsymbol{\mu}\|_{\mathbb{R}^n}(A).$$

Hint: *Use Proposition* **??**.

Section 5.4

Lebesgue measure on $\mathbb{R}$

In this section we specialise the general constructions of the preceding section to a special measure on the set $\mathbb{R}$. Our construction proceeds by first defining an outer measure, then using Theorem 5.3.13 to infer from this a complete measure space. The idea of measure that we use in this section is to be thought of as a generalisation of "length," and we shall point out as we go along that it does indeed share the features of "length" where the latter makes sense. However, the measure we define can be applied to sets for which it is perhaps not clear that a naïve definition of length is possible.

We shall see as we progress through this section that the σ-algebra we define is (1) not the collection of all subsets of $\mathbb{R}$ and (2) contains any reasonable set one could desire, and many more that one may not desire.

Do I need to read this section? If you are in the business of learning about the Lebesgue measure, this is where you go about it. •

5.4.1 The Lebesgue outer measure and the Lebesgue measure on $\mathbb{R}$

Our construction of the Lebesgue measure is carried out as per the idea in Section 5.3.2. That is to say, we construct an outer measure on $\mathbb{R}$ and take the measurable sets for this outer measure as the σ-algebra for the Lebesgue measure.

We first define the outer measure we use.

5.4.1 Definition (Lebesgue outer measure on $\mathbb{R}$) The ***Lebesgue outer measure*** on $\mathbb{R}$ is defined by

$$\lambda^*(S) = \inf\Big\{\sum_{j=1}^{\infty}|b_j - a_j| \;\Big|\; S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}}(a_j, b_j)\Big\}. \qquad \bullet$$

Thus the Lebesgue outer measure of $S \subseteq \mathbb{R}$ is the smallest sum of the lengths of open intervals that are needed to cover S. Let us define the length of a general interval I by

$$\ell(I) = \begin{cases} b - a, & \operatorname{cl}(I) = [a, b], \\ \infty, & I \text{ is unbounded.} \end{cases}$$

We next verify that the Lebesgue outer measure is indeed an outer measure, and we give its value on intervals.

5.4.2 Theorem (Lebesgue outer measure is an outer measure) *The Lebesgue outer measure is an outer measure on $\mathbb{R}$. Furthermore, if* I *is an interval then* $\lambda^*(\mathrm{I})$ *is the length of* I.

Proof First we show that $\lambda^*(\emptyset) = 0$. Indeed, let $(\epsilon_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence converging to zero in $\mathbb{R}_{>0}$ and note that $\emptyset \subseteq (-\epsilon_j, \epsilon_j)$, $j \in \mathbb{Z}_{>0}$. Since $\lim_{j\to\infty}|\epsilon_j + \epsilon_j| = 0$, our assertion follows.

Next we show that λ^* is monotonic. This is clear since if $A \subseteq B \subseteq \mathbb{R}$ and if a collection of intervals $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ covers B, then the same collection of intervals covers A.

For countable-subadditivity, let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a collection of subsets of $\mathbb{R}$. If $\sum_{j=1}^{\infty} \lambda^*(A_j) = \infty$ then countable-subadditivity follows trivially in this case, so we may as well suppose that $\sum_{j=1}^{\infty} \lambda^*(A_j) < \infty$. For $j \in \mathbb{Z}_{>0}$ and $\epsilon \in \mathbb{R}_{>0}$ let $((a_{j,k}, b_{j,k}))_{k\in\mathbb{Z}_{>0}}$ be a collection of open sets covering A_j and for which

$$\sum_{k=1}^{\infty} |b_{j,k} - a_{j,k}| < \lambda^*(A_j) + \frac{\epsilon}{2^j}.$$

By Proposition **??**, $\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ is countable. Therefore we may arrange the intervals $((a_{j,k}, b_{j,k}))_{j,k\in\mathbb{Z}_{>0}}$ into a single sequence $((a_n, b_n))_{n\in\mathbb{Z}_{>0}}$ so that

1. $\cup_{j\in\mathbb{Z}_{>0}} A_j \subseteq \cup_{n\in\mathbb{Z}_{>0}} (a_n, b_n)$ and
2. $\sum_{n=1}^{\infty} |b_n - a_n| < \sum_{n=1}^{\infty} \Big(\lambda^*(A_n) + \frac{\epsilon}{2^n}\Big) = \sum_{n=1}^{\infty} \lambda^*(A_n) + \epsilon.$

This shows that

$$\lambda^*\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) \le \sum_{n=1}^{\infty} \lambda^*(A_n),$$

giving countable-subadditivity.

We finally show that $\lambda^*(I) = \ell(I)$ for any interval I. We first take $I = [a, b]$. We may cover $[a, b]$ by $\{(a - \frac{\epsilon}{4}, b + \frac{\epsilon}{4})\} \cup ((0, \frac{\epsilon}{2^{j+1}}))_{j\in\mathbb{Z}_{>0}}$. Therefore,

$$\lambda^*([a, b]) \le (b + \tfrac{\epsilon}{4} - a + \tfrac{\epsilon}{4}) + \sum_{j=1}^{\infty} \frac{\epsilon}{2^{j+1}} = b - a + \epsilon,$$

where we use Example 2.4.2–**??**. Since ϵ can be made arbitrarily small we have $\lambda^*([a, b]) \le b - a$. Also, suppose that $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ covers $[a, b]$. By Theorem 2.5.27 there exists $n \in \mathbb{Z}_{>0}$ such that $[a, b] \subseteq \cup_{j=1}^{n} (a_j, b_j)$. Among the intervals $((a_j, b_j))_{j=1}^{n}$ we can pick a subset $((a_{j_k}, b_{j_k}))_{k=1}^{m}$ with the properties that $a \in (a_{j_1}, b_{j_1})$, $b \in (a_{j_m}, b_{j_m})$, and $b_{j_k} \in (a_{j_{k+1}}, b_{j_{k+1}})$. (Do this by choosing (a_{j_1}, b_{j_1}) such that a is in this interval. Then choose (a_{j_2}, b_{j_2}) such that b_{j_1} is in this interval. Since there are only finitely many intervals covering $[a, b]$, this can be continued and will stop by finding an interval containing b.) These intervals then clearly cover $[a, b]$ and also clearly satisfy $\sum_{k=1}^{m} |b_{j_k} - a_{j_k}| \ge b - a$ since they overlap. Thus we have

$$b - a \le \sum_{k=1}^{m} |b_{j_k} - a_{j_k}| \le \sum_{j=1}^{\infty} |b_j - a_j|.$$

Thus $b - a$ is a lower bound for the set

$$\Big\{\sum_{j=1}^{\infty} |b_j - a_j| \;\Big|\; [a, b] \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j)\Big\}.$$

Since $\lambda^*([a, b])$ is the greatest lower bound we have $\lambda^*([a, b]) \ge b-a$. Thus $\lambda^*([a, b]) = b-a$.

Now let I be a bounded interval and denote $\mathrm{cl}(I) = [a, b]$. Since $I \subseteq [a, b]$ we have $\lambda^*(I) \leq b - a$ using monotonicity of λ^*. If $\epsilon \in \mathbb{R}_{>0}$ we may find a closed interval $J \subseteq I$ for which the length of I exceeds that of J by at most ϵ. Since $\lambda^*(J) \leq \lambda^*(I)$ by monotonicity of λ^*, it follows that $\lambda^*(I)$ differs from the length of I by at most ϵ. Thus

$$\lambda^*(I) \geq \lambda^*(J) = b - a - \epsilon.$$

Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary $\lambda^*(I) \geq b - a$, showing that $\lambda^*(I) = b - a$, as desired.

Finally, if I is unbounded then for any $M \in \mathbb{R}_{>0}$ we may find a closed interval $J \subseteq I$ for which $\lambda^*(J) > M$. Since $\lambda^*(I) \geq \lambda^*(J)$ by monotonicity of λ^*, this means that $\lambda^*(I) = \infty$. ■

Now, having an outer measure on $\mathbb{R}$ one can ask, “Is λ^* a measure?” As we saw in Corollary 5.3.14 this amounts to asking, “Are all subsets of $\mathbb{R}$ λ^*-measurable?” Let us answer this question in the negative.

5.4.3 Example (A set that is not λ^*-measurable) Define an equivalence relation $\sim$ on $\mathbb{R}$ by

$$x \sim y \qquad \Longleftrightarrow \qquad x - y \in \mathbb{Q}.$$

By Proposition 1.2.9 it follows that $\mathbb{R}$ is the disjoint union of the equivalence classes for this equivalence relation. Moreover, each equivalence class has an element in the interval $(0, 1)$ since, for any $x \in \mathbb{R}$, the set

$$\{x + q \mid q \in \mathbb{Q}\}$$

intersects $(0, 1)$. By the Axiom of Choice, let $A \subseteq (0, 1)$ be defined by asking that A contain exactly one element from each equivalence class. We claim that A is not λ^*-measurable.

Let $\{q_j\}_{j \in \mathbb{Z}_{>0}}$ be an enumeration of the set of rational numbers in $(-1, 1)$ and for $j \in \mathbb{Z}_{>0}$ define

$$A_j = \{a + q_j \mid a \in A\}.$$

Note that $\cup_{j \in \mathbb{Z}_{>0}} A_j \subseteq (-1, 2)$.

We claim that $A_j \cap A_k \neq \emptyset$ if and only if $j = k$. Indeed, suppose that $A_j \cap A_k = \{x\}$. Then

$$x = a_j + q_j = a_k + q_k, \qquad a_j, a_k \in A.$$

Therefore, $a_j \sim a_k$ and, by construction of A, this implies that $a_j = a_k$. Thus $q_j = q_k$ and so $j = k$.

We also claim that $(0, 1) \subseteq \cup_{j \in \mathbb{Z}_{>0}} A_j$. Indeed, if $x \in (0, 1)$ then there exists $a \in A$ such that $x \sim a$. Note that $x - a \in \mathbb{Q} \cap (-1, 1)$ and so $x = a + q_j$ for some $j \in \mathbb{Z}_{>0}$. Thus $x \in A_j$.

Now suppose that A is λ^*-measurable. As we shall see in Theorem 5.4.23 below, this implies that A_j is λ^*-measurable for each $j \in \mathbb{Z}_{>0}$ and that $\lambda^*(A_j) = \lambda^*(A)$. We consider two cases.

1. $\lambda^*(A) = 0$: In this case, since the sets A_j, $j \in \mathbb{Z}_{>0}$, are disjoint, by properties of the measure we have

$$\mu\Big(\bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j) = 0.$$

But this contradicts the fact that $(0,1) \subseteq \cup_{j\in\mathbb{Z}_{>0}} A_j$.

2. $\lambda^*(A) \in \mathbb{R}_{>0}$: In this case we have

$$\mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j) = \infty.$$

But this contradicts the fact that $\cup_{j\in\mathbb{Z}_{>0}} A_j \subseteq (-1,2)$.

The contradiction that arises for both possibilities forces us to conclude that A is not measurable. •

Thus, making the following definition is not a vacuous procedure, and gives a strict subset of $\mathbf{2}^{\mathbb{R}}$ of λ^*-measurable sets.

5.4.4 Definition (Lebesgue measurable subset of $\mathbb{R}$, Lebesgue measure on $\mathbb{R}$) Let λ^* be the Lebesgue outer measure on $\mathbb{R}$ and denote by $\mathscr{L}(\mathbb{R})$ the set of λ^*-measurable subsets of $\mathbb{R}$. The sets in $\mathscr{L}(\mathbb{R})$ are called ***Lebesgue measurable***, or merely ***measurable***, and the complete measure $\lambda\colon \mathscr{L}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ induced by λ^* is called the ***Lebesgue measure*** on $\mathbb{R}$. •

The fairly concrete Example 5.4.3 can actually be sharpened considerably.

5.4.5 Theorem (The wealth of nonmeasurable subsets) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *satisfies* $\lambda(\mathrm{A}) \in \mathbb{R}_{>0}$ *then there exists* $\mathrm{S} \subseteq \mathrm{A}$ *that is not in* $\mathscr{L}(\mathbb{R})$.

Proof We have $A = \cup_{k\in\mathbb{Z}_{>0}}[-k,k] \cap A$ giving

$$0 < \lambda(A) \leq \sum_{k=1}^{\infty} \lambda([-k,k] \cap A).$$

Thus there exists $N \in \mathbb{Z}_{>0}$ such that $\lambda([-N,N] \cap A > 0$. Therefore, without loss of generality we may suppose that $A \subseteq [-N,N]$ for some $N \in \mathbb{Z}_{>0}$. Let $C \subseteq A$ be a countable subset of A and denote by H_C the subgroup of $(\mathbb{R},+)$ generated by C (see Definition 4.1.14). Therefore, by Proposition 4.1.13 it follows that

$$\mathsf{H}_C = \Big\{\sum_{j=1}^{k} n_j x_j \;\Big|\; k \in \mathbb{Z}_{>0},\ n_1,\dots,n_k \in \mathbb{Z}_{>0},\ x_1,\dots,x_k \in C\Big\}.$$

Note that H_C is then a countable union of countable sets and so is countable by Proposition **??**. Now note that the cosets of H_C form a partition of $\mathbb{R}$. Let $S' \subseteq \mathbb{R}$ be chosen (using the Axiom of Choice) such that S' contains exactly one representative from each coset of H_C. Then define

$$S = \{x \in A \mid x \in (x' + \mathsf{H}_C) \cap A,\ x' \in S'\}.$$

We will show that $S \notin \mathscr{L}(\mathbb{R})$.

For subsets $X, Y \subseteq \mathbb{R}$ let us denote

$$X + Y = \{x + y \mid x \in X,\ y \in Y\}, \quad X - Y = \{x - y \mid x \in X,\ y \in Y\}.$$

Let $B = \mathsf{H}_C \cap (A - A)$. Since $C - C \subseteq B$ we conclude that B is countable. We claim that if $(x_1 + S) \cap (x_2 + S) \neq \emptyset$ for $x_1, x_2 \in B$ then $x_1 = x_2$. Indeed, let $x \in (x_1 + S) \cap (x_2 + S)$ so that

$$x = x_1 + y_1 = x_2 + y_2, \qquad y_1, y_2 \in S.$$

Since $x_1, x_2 \in \mathsf{H}_C$ this implies that $y_2 - y_1 \in \mathsf{H}_C \cap S$ and so $y_1 = y_2$ by construction of S. Thus $(x + S)_{x \in B}$ is a family of pairwise disjoint sets. Moreover, $x + S \subseteq [-3N, 3N]$ for every $x \in B$ since $B, S \subseteq [-N, N]$. We further claim that $A \subseteq B + S$. Indeed, if $x \in A$ then x is in some coset of H_C: $x = y' + \mathsf{H}_C$ for $y' \in S'$. Then, since $x \in A$, there exists $y \in S$ such that $y + \mathsf{H}_C = y' + \mathsf{H}_C$. Thus $x = y + h$ for $y \in S$ and $h \in \mathsf{H}_C$. Therefore, $h = x - z \in A - A$ and so $h \in B$. Thus $x \in B + S$ as desired.

Now suppose that $S \in \mathscr{L}(\mathbb{R})$. There are two possibilities.

1. $\lambda(S) = 0$: In this case we have

$$\lambda(B + S) = \sum_{x \in B} \lambda(x + S) = \sum_{x \in B} \lambda(S) = 0,$$

where we have used the translation-invariance of the Lebesgue measure which we shall prove as Theorem 5.4.23 below. Since $A \subseteq B + S$ and $\lambda(A) \in \mathbb{R}_{>0}$ this is impossible.

2. $\lambda(S) \in \mathbb{R}_{>0}$: In this case we have

$$\lambda(B + S) = \sum_{x \in B} \lambda(x + S) = \sum_{x \in B} \lambda(S) = \infty.$$

Again, this is impossible, this time because $B + S \subseteq [-3N, 3N]$.

The impossibility of the two possible choices if S is Lebesgue measurable forces us to conclude that S is not Lebesgue measurable. ■

The reader might benefit by comparing the proof of the preceding theorem with the more concrete construction of Example 5.4.3.

We will very often wish to consider the Lebesgue measure not on all of $\mathbb{R}$, but on subsets of $\mathbb{R}$. Generally the subsets we consider will be intervals, but let us indicate how to restrict the Lebesgue measure to quite general subsets.

5.4.6 Proposition (Restriction of Lebesgue measure to measurable subsets) *Let* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *and denote*

(i) $\mathscr{L}(\mathrm{A}) = \{\mathrm{B} \cap \mathrm{A} \mid \mathrm{B} \in \mathscr{L}(\mathbb{R})\}$ *and*

(ii) $\lambda_{\mathrm{A}} \colon \mathscr{L}(\mathrm{A}) \to \overline{\mathbb{R}}_{\geq 0}$ *given by* $\lambda_{\mathrm{A}}(\mathrm{B} \cap \mathrm{A}) = \lambda(\mathrm{B} \cap \mathrm{A})$.

Then $(\mathrm{A}, \mathscr{L}(\mathrm{A}), \lambda_{\mathrm{A}})$ *is a complete measure space.*

Proof This follows from Propositions 5.2.6, 5.3.18, and 5.3.23. ■

5.4.2 Borel sets in $\mathbb{R}$ as examples of Lebesgue measurable sets

As we saw in Example 5.4.3, there are subsets of $\mathbb{R}$ that are not Lebesgue measurable. This then forces us to ask, "Which subsets of $\mathbb{R}$ *are* Lebesgue measurable?" To completely answer this question is rather difficult. What we shall do instead is provide a large collection of subsets that (1) are Lebesgue measurable, (2) are somewhat easy to understand (or at least convince ourselves that we understand), and (3) in an appropriate sense approximately characterise the Lebesgue measurable sets.

The sets we describe are given in the following definition. Denote by $\mathscr{O}(\mathbb{R}) \subseteq \mathbf{2}^{\mathbb{R}}$ be the collection of open subsets of $\mathbb{R}$.

5.4.7 Definition (Borel subsets of $\mathbb{R}$) The collection of ***Borel sets*** in $\mathbb{R}$ is the σ-algebra generated by $\mathscr{O}(\mathbb{R})$ (see Proposition 5.2.7). We denote by $\mathscr{B}(\mathbb{R})$ the Borel sets in $\mathbb{R}$. If $A \in \mathscr{B}(\mathbb{R})$ then we denote

$$\mathscr{B}(A) = \{A \cap B \mid B \in \mathscr{B}(\mathbb{R})\} \qquad \bullet$$

It is not so easy to provide a characterisation of the general Borel set, but certainly Borel sets can account for many sorts of sets. Borel sets are a large class of sets, and we shall pretty much only encounter Borel sets except when we are in the process of trying to be pathological. Furthermore, as we shall shortly see, Borel sets are Lebesgue measurable, and so serve to generate a large class of fairly easily described Lebesgue measurable sets.

Let us give some simple classes of Borel sets.

5.4.8 Examples (Borel sets)

1. All open sets are Borel sets, obviously.
2. All closed sets are Borel sets since closed sets are complements of open sets, and since σ-algebras are closed under complementation.
3. All intervals are Borel sets; Exercise 5.4.3.
4. The set $\mathbb{Q}$ of rational numbers is a Borel set; Exercise 5.4.4.
5. A subset $A \subseteq \mathbb{R}$ is a $\mathbf{G}_{\boldsymbol{\delta}}$ if $A = \cap_{j\in\mathbb{Z}_{>0}} O_j$ for a family $(O_j)_{j\in\mathbb{Z}_{>0}}$ of open sets. A G_δ is a Borel set; Exercise 5.4.5.
6. A subset $A \subseteq \mathbb{R}$ is an $\mathbf{F}_{\boldsymbol{\sigma}}$ if $A = \cup_{j\in\mathbb{Z}_{>0}} C_j$ for a family $(C_j)_{j\in\mathbb{Z}_{>0}}$ of closed sets. An F_σ is a Borel set; Exercise 5.4.5.

The practice of calling a set "a G_δ" or "an F_σ" is one of the unfortunate traditions involving poor notation in mathematics, notwithstanding that "G" stands for "Gebiet" ("open" in German), "F" stands for "fermé" ("closed" in French), "δ" stands for "Durchschnitt" ("intersection" in German), and "σ" stands for "Summe" ("sum" in German).

Let us first prove a result which gives interesting and sometimes useful alternative characterisations of Borel sets.

5.4.9 Proposition (Alternative characterisations of Borel sets) *$\mathscr{B}(\mathbb{R})$ is equal to the following collections of sets:*

(i) the σ-algebra $\mathscr{B}_1$ generated by the closed subsets;

(ii) the σ-algebra $\mathscr{B}_2$ generated by intervals of the form $(-\infty, \mathrm{b}]$, $\mathrm{b} \in \mathbb{R}$;

(iii) the σ-algebra $\mathscr{B}_3$ generated by intervals of the form $(\mathrm{a}, \mathrm{b}]$, $\mathrm{a}, \mathrm{b} \in \mathbb{R}$, $\mathrm{a} < \mathrm{b}$.

Proof First note that $\mathscr{B}(\mathbb{R})$ contains the σ-algebra $\mathscr{B}_1$ generated by all closed sets, since the complements of all open sets, i.e., all closed sets, are contained in $\mathscr{B}(\mathbb{R})$. Note that the sets of the form $(-\infty, b]$ are closed, so the σ-algebra $\mathscr{B}_2$ generated by these subsets is contained in $\mathscr{B}_1$. Since $(a, b] = (-\infty, b] \cap (\mathbb{R} \setminus (-\infty, a])$ it follows that the σ-algebra $\mathscr{B}_3$ generated by subsets of the form $(a, b]$ is contained in $\mathscr{B}_2$. Finally, note that

$$(a, b) = \cup_{n=1}^{\infty}(a, b - \tfrac{1}{n}].$$

Thus, by Proposition 2.5.6, it follows that every open set is a countable union of sets, each of which is a countable intersection of generators of $\mathscr{B}_3$. Thus $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{B}_3$. Putting this all together gives

$$\mathscr{B}(\mathbb{R}) \subseteq \mathscr{B}_3 \subseteq \mathscr{B}_2 \subseteq \mathscr{B}_1 \subseteq \mathscr{B}(\mathbb{R}).$$

Thus we must conclude that $\mathscr{B}_1 = \mathscr{B}_2 = \mathscr{B}_3 = \mathscr{B}(\mathbb{R})$. ■

We can then assert that all Borel sets are Lebesgue measurable.

5.4.10 Theorem (Borel sets are Lebesgue measurable) $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{L}(\mathbb{R})$.

Proof The theorem will follow from Proposition 5.4.9 if we can show that any set of the form $(-\infty, b]$ is Lebesgue measurable. Let A be such an interval and note that since

$$\lambda^*(S) \leq \lambda^*(S \cap A) + \lambda^*(S \cap (\mathbb{R} \setminus A))$$

we need only show the opposite inequality to show that A is Lebesgue measurable. If $\lambda^*(S) = \infty$ this is clearly true, so we may as well suppose that $\lambda^*(S) < \infty$. Let $((a_j, b_j))_{j \in \mathbb{Z}_{>0}}$ cover S so that

$$\sum_{j=1}^{\infty} |b_j - a_j| < \lambda^*(S) + \epsilon.$$

For $j \in \mathbb{Z}_{>0}$ choose intervals (c_j, d_j) and (e_j, f_j), possibly empty, for which

$$\begin{aligned}
&(a_j, b_j) \cap A \subseteq (c_j, d_j),\\
&(a_j, b_j) \cap (\mathbb{R} \setminus A) \subseteq (e_j, f_j),\\
&(d_j - c_j) + (f_j - e_j) \leq (b_j - a_j) + \frac{\epsilon}{2^j}.
\end{aligned}$$

Note that the intervals $((c_j, d_j))_{j \in \mathbb{Z}_{>0}}$ cover $S \cap A$ and that the intervals $((e_j, f_j))_{j \in \mathbb{Z}_{>0}}$ cover $\mathbb{R} \setminus A$ so that

$$\lambda^*(S \cap A) \leq \sum_{j=1}^{\infty} |d_j - c_j|, \quad \lambda^*(S \cap (\mathbb{R} \setminus A)) \leq \sum_{j=1}^{\infty} |f_j - e_j|.$$

From this we have

$$\lambda^*(S \cap A) + \lambda^*(S \cap (\mathbb{R} \setminus A)) \leq \sum_{j=1}^{\infty} |b_j - a_j| + \epsilon < \lambda^*(S) + 2\epsilon,$$

using the fact that $\sum_{j=1}^{\infty} \frac{1}{2^j} = 1$ by Example 2.4.2–??. Since ϵ can be taken arbitrarily small, the inequality

$$\lambda^*(S) \geq \lambda^*(S \cap A) + \lambda^*(S \cap (\mathbb{R} \setminus A))$$

follows, and so too does the result. ■

The next result sharpens the preceding assertion considerably.

5.4.11 Theorem (Lebesgue measurable sets are the completion of the Borel sets) $(\mathbb{R}, \mathscr{L}(\mathbb{R}), \lambda)$ *is the completion of* $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda|\mathscr{B}(\mathbb{R}))$.

Proof First, given $A \in \mathscr{L}(\mathbb{R})$, we find $L, U \in \mathscr{B}(\mathbb{R})$ such that $L \subseteq A \subseteq U$ and such that $\lambda(U \setminus L) = 0$. We first suppose that $\lambda(A) < \infty$. Using Theorem 5.4.19 below, let $(U_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of open sets containing A and for which $\lambda(U_j) \leq \lambda(A) + \frac{1}{j}$ and let $(L_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of compact subsets of A for which $\lambda(L_j) \geq \lambda(A) - \frac{1}{j}$. If we take $L = \cup_{j\in\mathbb{Z}_{>0}} L_j$ and $U = \cap_{j\in\mathbb{Z}_{>0}} U_j$ then we have $L \subseteq A \subseteq U$. We also have

$$\lambda(U \setminus L) \leq \lambda(U_j \setminus L_j) = \lambda(U_j \setminus A) + \lambda(A \setminus L_j) \leq \tfrac{1}{2j}.$$

Since this holds for every $j \in \mathbb{Z}_{>0}$, this gives our claim when A has finite measure, since L and U are Borel sets. If $\lambda(A) = \infty$ then we can write $A = \cup_{j\in\mathbb{Z}_{>0}} A_j$ with $A_j = (-j, j) \cap A$. For each $j \in \mathbb{Z}_{>0}$ we may find $L_j, U_j \in \mathscr{B}(\mathbb{R})$ such that $L_j \subseteq A_j \subseteq U_j$ and $\lambda(U_j \setminus L_j)$. Taking $L = \cup_{j\in\mathbb{Z}_{>0}} L_j$ and $U = \cup_{j\in\mathbb{Z}_{>0}}$ gives $L \subseteq A \subseteq U$ and $\lambda(U \setminus L) = 0$.

The above shows that $\mathscr{L}(\mathbb{R}) \subseteq \mathscr{B}_\lambda(\mathbb{R})$. Now let $B \in \mathscr{B}_\lambda(\mathbb{R})$ and take Borel sets L and U for which $L \subseteq B \subseteq U$ and $\lambda(U \setminus L) = 0$. Note that $(B \setminus L) \subseteq (U \setminus L)$. Note also that since $U \setminus L \in \mathscr{B}(\mathbb{R})$ we have $U \setminus L \in \mathscr{L}(\mathbb{R})$ and $\lambda(U \setminus L) = 0$. By completeness of the Lebesgue measure this implies that $B \setminus L \in \mathscr{L}(\mathbb{R})$. Since $B = (B \setminus L) \cup L$ this implies that $B \in \mathscr{L}(\mathbb{R})$. ■

The following corollary indicates that Borel sets closely approximate Lebesgue measurable sets.

5.4.12 Corollary (Borel approximations to Lebesgue measurable sets) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *then there exists a Borel set* B *and a set* Z *of measure zero such that* $\mathrm{A} = \mathrm{B} \cup \mathrm{Z}$.

Proof This follows directly from Theorem 5.4.11 and the definition of the completion. ■

The preceding result looks like good news in that, except for seemingly irrelevant sets of measure zero, Lebesgue measurable sets agree with Borel sets. The problem is that there are lots of sets of measure zero. The following result indicates that this is reflected by a big difference in the number of Lebesgue measurable sets versus the number of Borel sets.

5.4.13 Proposition (The cardinalities of Borel and Lebesgue measurable sets) *We have* $\mathrm{card}(\mathscr{B}(\mathbb{R})) = \mathrm{card}(\mathbb{R})$ *and* $\mathrm{card}(\mathscr{L}(\mathbb{R})) = \mathrm{card}(\mathbf{2}^{\mathbb{R}})$.

Proof Since $\{x\} \in \mathscr{B}(\mathbb{R})$ for every $x \in \mathbb{R}$ is follows that $\mathrm{card}(\mathscr{B}(\mathbb{R})) \geq \mathrm{card}(\mathbb{R})$. Let $\mathscr{O}_\mathbb{Q}$ be the collection of open intervals with rational (or infinite) endpoints. The set $\mathscr{O}_\mathbb{Q}$ is a countable union of countable sets and so is countable by Proposition **??**. Since every open set is a countable union of sets from $\mathscr{O}_\mathbb{Q}$ (cf. Proposition 2.5.6 and see Proposition **??**) it that if we take $\mathscr{S} = \mathscr{O}_\mathbb{Q}$ then, in the notation of Theorem 5.2.14, $\mathscr{S}_1$ includes the collection of open sets. Then it follows that $\mathscr{B}(\mathbb{R})$ is the σ-algebra generated by the countable family $\mathscr{O}_\mathbb{Q}$ of subsets of $\mathbb{R}$. By Theorem 5.2.14 it follows that $\mathrm{card}(\mathscr{B}(\mathbb{R})) \leq \aleph_0^{\aleph_0} = \mathrm{card}(\mathbb{R})$, using the computation

$$2^{\aleph_0} \leq \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0},$$

which holds since $2 \leq \aleph_0 \leq 2^{\aleph_0}$ by Example **??**–**??** and Exercise **??**.

To show that $\mathrm{card}(\mathscr{L}(\mathbb{R})) = \mathrm{card}(\mathbf{2}^{\mathbb{R}})$ first note that $\mathrm{card}(\mathscr{L}(\mathbb{R})) \leq \mathrm{card}(\mathbf{2}^{\mathbb{R}})$. For the opposite inequality, recall from Example 2.5.39 that the middle-thirds Cantor set $C \subseteq [0,1]$ has the properties (1) $\lambda(C) = 0$ and (2) $\mathrm{card}(C) = \mathrm{card}([0,1]) = \mathrm{card}(\mathbb{R})$. Since the Lebesgue measure is complete, every subset of C is Lebesgue measurable and has Lebesgue measure zero. This shows that $\mathrm{card}(\mathbf{2}^{C}) = \mathrm{card}(\mathbf{2}^{\mathbb{R}}) \leq \mathrm{card}(\mathscr{L}(\mathbb{R}))$. ■

While the preceding result is interesting in that it tells us that there are many more Lebesgue measurable sets than Borel sets, Corollary 5.4.12 notwithstanding, it does not tell us what a non-Borel Lebesgue measurable set might look like. The following is a concrete example of such a set. Our construction uses some facts about measurable functions that we will not introduce until Section 5.6.

5.4.14 Example (A non-Borel Lebesgue measurable set) Recall from Example 3.2.27 the construction of the Cantor function $f_C\colon [0,1] \to [0,1]$, and recall that f_C is continuous, monotonically increasing, and satisfies $f_C(0) = 0$ and $f_C(1) = 1$. Thus, by the Intermediate Value Theorem, for each $y \in [0,1]$ there exists $x \in [0,1]$ such that $f_C(x) = y$. We use this fact to define $g_C\colon [0,1] \to [0,1]$ by

$$g_C(y) = \inf\{x \in [0,1] \mid f_C(x) = y\}.$$

Let us prove some facts about g_C.

1 Lemma *We have* $f_C \circ g_C(y) = y$ *and so* g_C *is injective.*

Proof Let $(x_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $[0,1]$ for which $\lim_{j\to\infty} f_C(x_j) = y$. This sequence contains a convergent subsequence $(x_{j_k})_{k\in\mathbb{Z}_{>0}}$ by the Bolzano–Weierstrass Theorem; let $x = \lim_{k\to\infty} x_{j_k}$. Then, by continuity of f_C, $y = f_C(x)$. We also have $g_C(y) = x$ by definition, and so this gives $f_C \circ g_C(y) = y$, as desired. Injectivity of g_C follows from Proposition 1.3.9. ▼

2 Lemma *The function* g_C *is monotonically increasing.*

Proof Let $y_1, y_2 \in [0,1]$ satisfy $y_1 < y_2$ and suppose that $g_C(y_1) > g_C(y_2)$. Then $f_C \circ g_C(y_1) \geq f_C \circ g_C(y_2)$ since f_C is monotonically increasing. From the previous lemma this implies that $y_1 \geq y_2$ which is a contradiction. Thus we must have $g_C(y_1) \leq g_C(y_2)$. ▼

3 Lemma $\mathrm{image}(g_C) \subseteq C$.

Proof For $y \in \mathbb{R}$ the set

$$\{x \in [0,1] \mid f_C(x) = y\}$$

is an interval, possibly with empty interior, on which f_C is constant. The endpoints of the interval are points in C. In particular, $g_C(y) \in C$. ▼

Now let $A \subseteq [0,1]$ be the non-Lebesgue measurable subset of Example 5.4.3 and take $B = g_C(A)$. Then $B \subseteq C$ and so is a subset of a set of measure zero by Example 2.5.39. Since the Lebesgue measure is complete it follows that B is Lebesgue measurable. However, were B to be a Borel set, then monotonicity of g_C

and ***missing stuff*** implies that $g_C^{-1}(B)$ is a Borel set. However, injectivity of g_C gives $g_C^{-1}(B) = A$, and A is not Lebesgue measurable, and so certainly not Borel. Thus B is not a Borel set. •

When we come to talk about functions defined on measurable spaces in Section 5.6 we will consider functions taking values in $\overline{\mathbb{R}}$. It will then be occasionally useful to have a notion of a Borel subset of $\overline{\mathbb{R}}$. Let us, therefore, define what these subsets are.

5.4.15 Definition (Borel subsets of $\overline{\mathbb{R}}$) The collection of ***Borel sets*** in $\overline{\mathbb{R}}$ is the σ-algebra generated by the subsets of $\overline{\mathbb{R}}$ having the following form:

$$U, \quad U \cup [-\infty, b), \quad U \cup (a, \infty], \quad U \cup [-\infty, b) \cup (a, \infty], \qquad U \in \mathscr{O}(\mathbb{R}),\ a, b \in \mathbb{R}.$$

We denote by $\mathscr{B}(\overline{\mathbb{R}})$ the Borel sets in $\overline{\mathbb{R}}$. •

The idea of the preceding definition is that $\mathscr{B}(\overline{\mathbb{R}})$ is the σ-algebra generated by open subsets of $\overline{\mathbb{R}}$, where open subsets of $\overline{\mathbb{R}}$ are those used in the definition. That these open subsets are indeed the open subsets for a topology on $\overline{\mathbb{R}}$ is argued in Example **??–??**.

The following characterisation of $\mathscr{B}(\overline{\mathbb{R}})$ is useful.

5.4.16 Proposition (Characterisation of $\mathscr{B}(\overline{\mathbb{R}})$) *The σ-algebra $\mathscr{B}(\overline{\mathbb{R}})$ is generated by $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$.*

Proof Clearly $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{B}(\overline{\mathbb{R}})$. Since

$$\{\infty\} = \cap_{k \in \mathbb{Z}}(k, \infty], \quad \{-\infty\} = \cap_{k \in \mathbb{Z}}[-\infty, -k),$$

and since $(k, \infty], [-\infty, -k) \in \mathscr{B}(\overline{\mathbb{R}})$ for each $k \in \mathbb{Z}_{>0}$, it follows that $\{-\infty\}, \{\infty\} \in \mathscr{B}(\overline{\mathbb{R}})$. Therefore, the σ-algebra generated by $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$ is contained in $\mathscr{B}(\overline{\mathbb{R}})$.

Next we note that $U \in \mathscr{B}(\mathbb{R})$ if $U \in \mathscr{O}(\mathbb{R})$. Also, for $b \in \mathbb{R}$,

$$U \cup [-\infty, b) = U \cup \{-\infty\} \cup (-\infty, b)$$

and so $U \cup [-\infty, b)$ is a union of sets from $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$. In similar fashion sets of the form

$$U \cup (a, \infty], \quad U \cup [-\infty, b) \cup (a, \infty]$$

for $a, b \in \mathbb{R}$ are unions of sets from $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$. This implies that the generators for the σ-algebra $\mathscr{B}(\overline{\mathbb{R}})$ are contained in the σ-algebra generated by $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$. Thus $\mathscr{B}(\overline{\mathbb{R}})$ is contained in the σ-algebra generated by $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$. ■

5.4.3 Further properties of the Lebesgue measure on $\mathbb{R}$

In this section we give some additional properties of the Lebesgue measure that (1) illustrate a sort of friendliness of this measure and (2) justify its being in some way natural.

Let us illustrate first an important property of the Lebesgue measure. Let us do this by giving a general definition that creates a little context for this property of Lebesgue measure.

5.4.17 Definition (Regular measure on $\mathbb{R}$) Let $\mathscr{A}$ be a σ-algebra on $\mathbb{R}$ that contains the Borel σ-algebra $\mathscr{B}(\mathbb{R})$. A measure $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ is ***regular*** if

(i) $\mu(K) < \infty$ for each compact subset $K \subseteq \mathbb{R}$,

(ii) if $A \in \mathscr{A}$ then $\mu(A) = \inf\{\mu(U) \mid U \text{ open and } A \subseteq U\}$, and

(iii) if $U \subseteq \mathbb{R}$ is open then $\mu(U) = \sup\{\mu(K) \mid K \text{ open and } K \subseteq U\}$. •

Before we prove that the Lebesgue measure is regular, let us give some examples that show that irregular measures are possible.

5.4.18 Examples (Regular and irregular measures)

1. For $x \in \mathbb{R}$ the point mass measure $\delta_x\colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$\delta(B) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B \end{cases}$$

is regular, as may be readily verified; see Exercise 5.4.6.

2. One can check that the counting measure $\mu\colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$\mu(B) = \begin{cases} \operatorname{card}(B), & \operatorname{card}(B) < \infty, \\ \infty, & \text{otherwise} \end{cases}$$

is not regular; see Exercise 5.4.7. •

We begin with a theorem that characterises the Lebesgue measure of measurable sets.

5.4.19 Theorem (Regularity of the Lebesgue measure) *The Lebesgue measure $\lambda\colon \mathscr{L}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ is σ-finite and regular. Moreover, for $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ we have $\lambda(\mathrm{A}) = \sup\{\lambda(\mathrm{K}) \mid K \text{ compact and } K \subseteq A\}$.*

Proof To see that λ is σ-finite note that $\mathbb{R} = \cup_{k\in\mathbb{Z}_{>0}}[-k,k]$ with $\lambda([-k,k]) < \infty$.

Next we show that if $A \in \mathscr{L}(\mathbb{R})$ then

$$\lambda(A) = \inf\{\lambda(U) \mid U \text{ open and } A \subseteq U\}.$$

Assume that $\lambda(A) < \infty$ since the result is obvious otherwise. Let $\epsilon \in \mathbb{R}_{>0}$ and let $((a_j,b_j))_{j\in\mathbb{Z}_{>0}}$ be a sequence of open intervals for which $A \subseteq \cup_{j\in\mathbb{Z}_{>0}}(a_j,b_j)$ and for which

$$\sum_{j=1}^{\infty}|b_j - a_j| = \lambda(A) + \epsilon.$$

Now let $U = \cup_{j\in\mathbb{Z}_{>0}}(a_j,b_j)$, noting that U is open and that $A \subseteq U$. By Proposition 5.3.10(iii) and the fact that the measure of an interval is its length we have

$$\lambda(U) \leq \sum_{j=1}^{\infty}|b_j - a_j| = \lambda(A) + \epsilon.$$

Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary this shows that

$$\lambda(A) \geq \inf\{\lambda(U) \mid U \text{ open and } A \subseteq U\}.$$

Since the other inequality is obvious by the basic properties of a measure, this part of the result follows.

Note that to show that λ is regular it suffices to prove the final assertion of the theorem since open sets are Lebesgue measurable; thus we prove the final assertion of the theorem. First suppose that $A \in \mathscr{L}(\mathbb{R})$ is bounded. Then let $\tilde{K}$ be a compact set containing A. For $\epsilon \in \mathbb{R}_{>0}$ choose U open and containing $\tilde{K} \setminus A$ and for which $\lambda(U) \leq \lambda(\tilde{K} \setminus A) + \epsilon$, this being possible from by the first part of the proof. Note that $K = \tilde{K} \setminus U$ is then a compact set contained in A and that the basic properties of measure then give

$$\lambda(U) \leq \lambda(\tilde{K} \setminus A) + \epsilon \text{ and } \lambda(\tilde{K}) \leq \lambda(K) - \lambda(A) \quad \Longrightarrow \quad \lambda(K) > \lambda(A) - \epsilon.$$

Since ϵ can be made as small as desired, this gives the second part of the proposition when A is bounded. Define

$$A_j = (-j, j) \cap A,$$

and note that $(A_j)_{j\in\mathbb{Z}_{>0}}$ is an increasing sequence of sets and that $A = \cup_{j\in\mathbb{Z}_{>0}} A_j$. Therefore, by Proposition 5.3.10(iv), $\lambda(A) = \lim_{j\to\infty} \lambda(A_j)$. Then for any $M < \lambda(A)$ there exists $N \in \mathbb{Z}_{>0}$ such that $\lambda(A_N) > M$. We may now find a compact K such that $\lambda(K) > M$ by the fact that we have proved our assertion for bounded sets (as is A_N). Note that $K \subseteq A$ and that $M < \lambda(A)$ is arbitrary, and so the result follows. ■

This result has the following obvious corollary.

5.4.20 Corollary (Approximation of Lebesgue measurable sets by open and compact sets) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *satisfies* $\lambda(\mathrm{A}) < \infty$ *and if* $\epsilon \in \mathbb{R}_{>0}$ *then there exists an open set* $\mathrm{U} \subseteq \mathbb{R}$ *and a compact set* $\mathrm{K} \subseteq \mathbb{R}$ *such that*

$$\lambda(\mathrm{U} \setminus \mathrm{A}) < \epsilon, \quad \lambda(\mathrm{A} \setminus \mathrm{K}) < \epsilon.$$

Let us next show that the Lebesgue measure is, in some way, natural. We do this by considering a particular property of the Lebesgue measure, namely that it is “translation-invariant.” In order to define what it means for a measure to be translation-invariant, we first need to say what it means for a σ-algebra to be translation-invariant.

5.4.21 Definition (Translation-invariant σ-algebra and measure on $\mathbb{R}$) A σ-algebra $\mathscr{A} \subseteq 2^{\mathbb{R}}$ is ***translation-invariant*** if, for every $A \in \mathscr{A}$ and every $x \in \mathbb{R}$,

$$x + A \triangleq \{x + y \mid y \in A\} \in \mathscr{A}.$$

A ***translation-invariant*** measure on a translation-invariant σ-algebra $\mathscr{A}$ is a map $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ for which $\mu(x + A) = \mu(A)$ for every $A \in \mathscr{A}$ and $x \in \mathbb{R}$. •

The two σ-algebras we are considering in this section are translation-invariant.

5.4.22 Proposition (Translation-invariance of Borel and Lebesgue measurable sets) *Both $\mathscr{B}(\mathbb{R})$ and $\mathscr{L}(\mathbb{R})$ are translation-invariant.*

Proof Let us denote

$$\mathscr{B}'(\mathbb{R}) = \{B \mid x + B \in \mathscr{B}(\mathbb{R}) \text{ for every } x \in \mathbb{R}\}.$$

We claim that $\mathscr{B}'(\mathbb{R})$ is a σ-algebra containing the open subsets of $\mathbb{R}$. First of all, if $U \subseteq \mathbb{R}$ is open then $x + U$ is open for every $x \in \mathbb{R}$ (why?) and so $U \in \mathscr{B}'(\mathbb{R})$. To see that $\mathscr{B}'(\mathbb{R})$ is a σ-algebra, first note that $\mathbb{R} = x + \mathbb{R}$ for every $x \in \mathbb{R}$ and so $\mathbb{R} \in \mathscr{B}'(\mathbb{R})$. Next, let $B \in \mathscr{B}'(\mathbb{R})$ and let $x \in \mathbb{R}$. Then

$$\begin{aligned} x + (\mathbb{R} \setminus B) &= \{x + z \mid z \notin B\} = \{y \mid y - x \notin B\} = \{y \mid y \neq x + z,\ z \in B\} \\ &= \{y \mid y \notin (x + B)\} = \mathbb{R} \setminus (x + B) \in \mathscr{B}(\mathbb{R}). \end{aligned}$$

Thus $x + (\mathbb{R} \setminus B) \in \mathscr{B}(\mathbb{R})$ for every $x \in \mathbb{R}$ and so $\mathbb{R} \setminus B \in \mathscr{B}'(\mathbb{R})$. Finally, let $(B_j)_{j\mathbb{Z}_{>0}}$ be a countable collection of subsets from $\mathscr{B}'(\mathbb{R})$. Then, for $x \in \mathbb{R}$ we have

$$x + \cup_{j\in\mathbb{Z}_{>0}} B_j = \cup_{j\in\mathbb{Z}_{>0}} (x + B_j) \in \mathscr{B}(\mathbb{R})$$

and so $\cup_{j\in\mathbb{Z}_{>0}} B_j \in \mathscr{B}'(\mathbb{R})$. Thus $\mathscr{B}'(\mathbb{R})$ is indeed a σ-algebra containing the open sets and so we conclude that $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{B}'(\mathbb{R})$ since $\mathscr{B}(\mathbb{R})$ is the σ-algebra generated by the open sets. This shows that $\mathscr{B}(\mathbb{R})$ is translation-invariant.

Next let us show that $\mathscr{L}(\mathbb{R})$ is translation-invariant. To do this we first show that if $S \subseteq \mathbb{R}$ and if $x \in \mathbb{R}$ then $\lambda^*(x + S) = \lambda^*(S)$. Indeed,

$$\begin{aligned} \lambda^*(x + S) &= \inf\Big\{ \sum_{j=1}^{\infty} |b_j - a_j| \;\Big|\; x + S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j) \Big\} \\ &= \inf\Big\{ \sum_{j=1}^{\infty} |b_j - a_j| \;\Big|\; x + S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (x + a_j, x + b_j) \Big\} \\ &= \inf\Big\{ \sum_{j=1}^{\infty} |b_j - a_j| \;\Big|\; S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} (a_j, b_j) \Big\} = \lambda^*(S). \end{aligned}$$

Now let $A \in \mathscr{L}(\mathbb{R})$ so that, for every subset $S \subseteq \mathbb{R}$,

$$\lambda^*(S) = \lambda^*(S \cap A) + \lambda^*(S \cap (\mathbb{R} \setminus A)).$$

Then, for $x \in \mathbb{R}$ and $S \subseteq \mathbb{R}$,

$$\lambda^*(S \cap (x + A)) = \lambda^*((x + (-x + S)) \cap (x + A)) = \lambda^*((-x + S) \cap A)$$

and, similarly,

$$\lambda^*(S \cap (\mathbb{R} \setminus (x + A))) = \lambda^*((x + (-x + S)) \cap (x + \mathbb{R} \setminus A)) = \lambda^*((-x + S) \cap (\mathbb{R} \setminus A)).$$

Since $\lambda^*(-x + S) = \lambda^*(S)$ this immediately gives

$$\lambda^*(S) = \lambda^*(S \cap (x + A)) + \lambda^*(S \cap (\mathbb{R} \setminus (x + A))),$$

showing that $x + A \in \mathscr{L}(\mathbb{R})$. ■

Now that the σ-algebras are known to be translation-invariant, we can make the following characterisation of the Lebesgue measure.

5.4.23 Theorem (Translation invariance of the Lebesgue measure) *If $\mu\colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ is a nonzero translation-invariant measure for which $\mu(\mathrm{B}) < \infty$ for every bounded $\mathrm{B} \in \mathscr{B}(\mathbb{R})$, then there exists $\mathrm{c} \in \mathbb{R}_{>0}$ such that $\mu(\mathrm{B}) = \mathrm{c}\lambda(\mathrm{B})$ for every $\mathrm{B} \in \mathscr{B}(\mathbb{R})$. Moreover, the Lebesgue measure $\lambda\colon \mathscr{L}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ is translation-invariant.*

Proof That λ is translation-invariant follows from the proof of Proposition 5.4.22 where we showed that $\lambda^*(x + S) = \lambda^*(S)$ for every $S \subseteq \mathbb{R}$ and $x \in \mathbb{R}$. To show that λ is, up to a positive scalar, the only translation-invariant measure we first prove two lemmata.

1 Lemma *If $\mathrm{U} \subseteq \mathbb{R}$ is a nonempty open set, then there exists a countable collection of disjoint half-open intervals $(\mathrm{I}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ such that $\mathrm{U} = \cup_{\mathrm{j}\in\mathbb{Z}_{>0}}\mathrm{I}_\mathrm{j}$.*

Proof For $k \in \mathbb{Z}_{\geq 0}$ define

$$\mathscr{C}_k = \{[j2^{-k}, (j+1)2^{-k}) \mid j \in \mathbb{Z}\}.$$

Note that, for each $k \in \mathbb{Z}_{\geq 0}$, the sets from $\mathscr{C}_k$ form a countable partition of $\mathbb{R}$. Also note that for $k < l$, every interval in $\mathscr{C}_l$ is also an interval in $\mathscr{C}_k$. Now let $U \subseteq \mathbb{R}$ be open. Let $\mathscr{D}_0 = \emptyset$. Let

$$\begin{aligned}
\mathscr{D}_1 &= \{I \in \mathscr{C}_1 \mid I \subseteq U\},\\
\mathscr{D}_2 &= \{I \in \mathscr{C}_2 \mid I \subseteq U,\ I \notin \mathscr{D}_1\},\\
&\ \ \vdots\\
\mathscr{D}_k &= \{I \in \mathscr{C}_k \mid I \subseteq U,\ I \notin \mathscr{D}_1 \cup \cdots \cup \mathscr{D}_{k-1}\}\\
&\ \ \vdots
\end{aligned}$$

The result will follow if we can show that each point $x \in U$ is contained in some $\mathscr{D}_k$, $k \in \mathbb{Z}_{>0}$. However, this follows since U is open, and so, for each $x \in U$, one can find a smallest $k \in \mathbb{Z}_{\geq 0}$ with the property that there exists $I \in \mathscr{C}_k$ with $x \in I$ and $I \subseteq U$. ▼

2 Lemma *The Lebesgue measure is the unique measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ for which the measure of an interval is its length.*

Proof From Theorem 5.4.2 we know that $\lambda(I) = \ell(I)$ for every interval I. Now suppose that $\mu\colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ is a measure with the property that $\mu(I) = \ell(I)$ for every interval I.

First let $U \subseteq \mathbb{R}$ be open. By Lemma 1 we can write $U = \cup_{j\in\mathbb{Z}_{>0}} I_j$ for a countable family $(I_j)_{j\in\mathbb{Z}_{>0}}$ of disjoint intervals. Therefore, since μ is a measure,

$$\mu(U) = \mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} I_j\Big) = \sum_{j=1}^{\infty} \mu(I_j) = \sum_{j=1}^{\infty} \lambda(I_j) = \lambda(U).$$

Now let B be a bounded Borel set and let U be an open set for which $B \subseteq U$. Then

$$\mu(B) \leq \mu(U) = \lambda(U).$$

Therefore,

$$\mu(B) \leq \inf\{\lambda(U) \mid U \text{ open and } B \subseteq U\} = \lambda(B)$$

by regularity of λ. Therefore, if U is a bounded open set containing B we have

$$\mu(U) = \mu(B) + \mu(U \setminus B) \leq \lambda(B) + \lambda(U \setminus B) = \lambda(U).$$

Since $\mu(U) = \lambda(U)$ it follows that $\mu(B) = \lambda(B)$ and $\mu(U \setminus B) = \lambda(U \setminus B)$.

Finally let B be an unbounded Borel set. We can then write $B = \cup_{j\in\mathbb{Z}} B_j$ where $(B_j)_{j\in\mathbb{Z}_{>0}}$ is the countable family of disjoint Borel sets $B_j = B \cap [j, j+1)$, $j \in \mathbb{Z}$. Then

$$\mu(B) = \sum_{j\in\mathbb{Z}} \mu(B_j) = \sum_{j\in\mathbb{Z}} \lambda(B_j) = \lambda(B),$$

as desired. ▼

To proceed with the proof, let $\mu\colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ be a translation-invariant measure and let $c = \mu([0,1))$. By assumption $c \in \mathbb{R}_{>0}$ since, were $c = 0$,

$$\mu(\mathbb{R}) = \sum_{j=1}^{\infty} \mu([j, j+1)) = 0$$

by translation-invariance of μ. Now let $\mu'\colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ be the measure defined by $\mu'(B) = c^{-1}\mu(B)$. Now, for $k \in \mathbb{Z}_{\geq 0}$ let

$$\mathscr{C}_k = \{[j2^{-k}, (j+1)2^{-k}) \mid j \in \mathbb{Z}\}$$

as in the proof of Lemma 1. Let $I \in \mathscr{C}_k$. We can write $[0,1)$ as a disjoint union of 2^k intervals of the form $x_j + I$. Therefore, by translation-invariance of μ',

$$\mu'([0,1)) = 2^k \mu'(I), \quad \lambda([0,1)) = 2^k \lambda(I).$$

Since $\mu'([0,1)) = \lambda([0,1))$ it follows that $\mu'(I) = \lambda(I)$. Since every interval is a disjoint union of intervals from the sets $\mathscr{C}_k$, $k \in \mathbb{Z}_{\geq 0}$, by Lemma 1 it follows that $\mu'(I) = \lambda(I)$ for every interval I. Thus $\mu' = \lambda$ by Lemma 2 above and so $\mu = c\lambda$, as desired. ■

It is then natural question to ask, "Are there larger σ-algebras than $\mathscr{B}(\mathbb{R})$ which admit a translation-invariant measure?" Obviously one such is the collection $\mathscr{L}(\mathbb{R})$ of Lebesgue measurable sets. But are there larger ones? The following result gives a partial answer, and indicates that the "best possible" construction is impossible.

5.4.24 Theorem (There are no translation-invariant, length-preserving measures on all subsets of $\mathbb{R}$) *There exists no measure space $(\mathbb{R}, \mathscr{A}, \mu)$ having the joint properties that*

(i) $\mathscr{A} = 2^{\mathbb{R}}$,

(ii) $\mu((0,1)) = 1$, *and*

(iii) μ *is translation-invariant.*

Proof Were such a measure to exist, then the non-Lebesgue measurable set $A \subseteq (0,1)$ of Example 5.4.3 would be measurable. But during the course of Example 5.4.3 we showed that $(0,1)$ is a countable disjoint union of translates of A. The dichotomy illustrated in Example 5.4.3 then applies. That is, if $\mu(A) = 0$ then we get $\mu((0,1)) = 0$ and if $\mu(A) \in \mathbb{R}_{>0}$ then $\mu((0,1)) = \infty$, both of which conclusions are false. ■

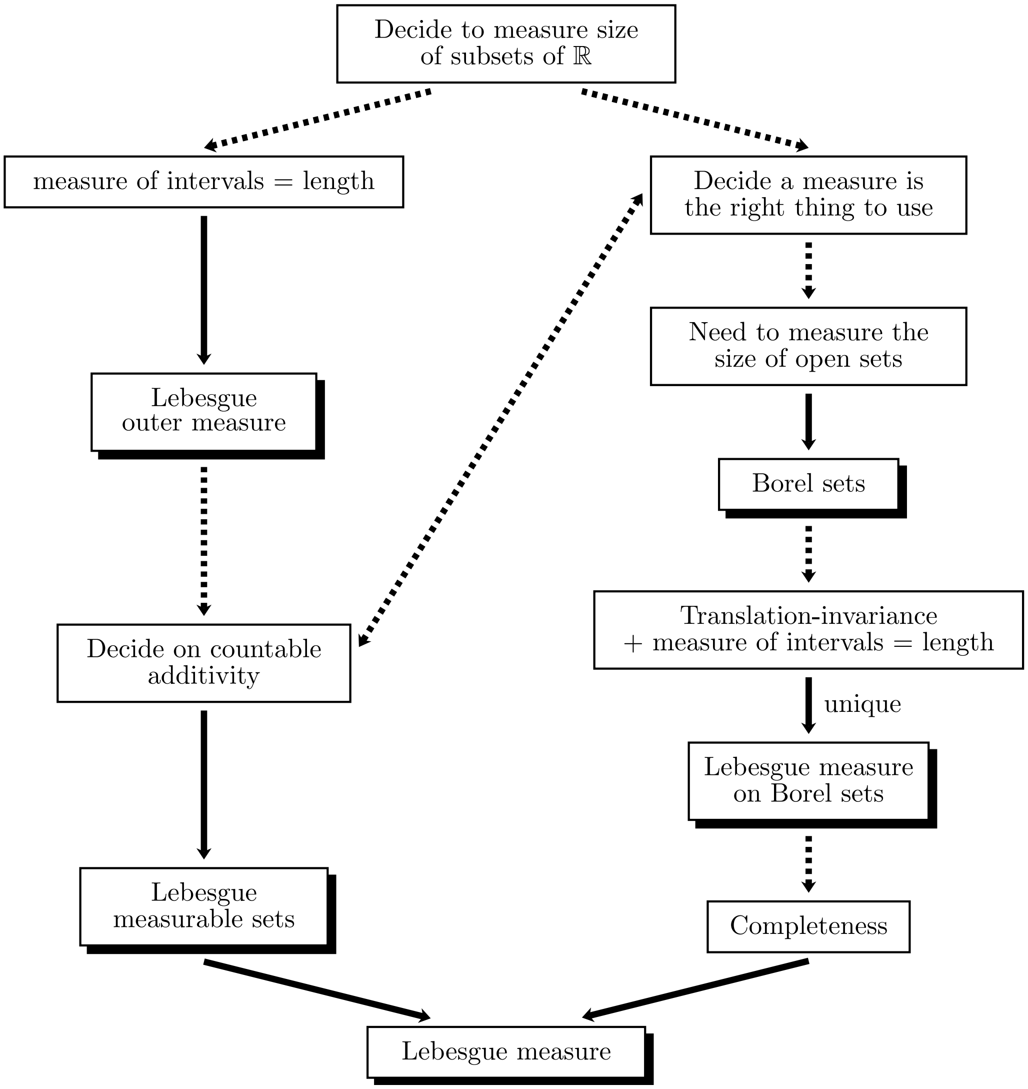

Decide to measure size of subsets of $\mathbb{R}$
measure of intervals = length
Decide a measure is the right thing to use
Need to measure the size of open sets
Lebesgue outer measure
Borel sets
Translation-invariance + measure of intervals = length
Decide on countable additivity
unique
Lebesgue measure on Borel sets
Lebesgue measurable sets
Completeness
Lebesgue measure

Figure 5.3 Lines of reasoning for arriving at Lebesgue measure. Dashed arrows represent choices that can be made and solid arrows represent conclusions that follow from the preceding decisions

It is now possible to provide a summary of the “reasonableness” of the Lebesgue measure by providing a natural line of reasoning, the natural terminus of which is the Lebesgue measure. In Figure 5.3 we show a “flow chart” for how one might justify the Lebesgue measure as being the process of some rational line of thought. Note that we are not saying that this actually described the historical development of the Lebesgue measure, but just that, after the fact, it indicates that the Lebesgue measure is not a strange thing to arrive at. It is rare that scientific discovery actually proceeds along the lines that make it most understandable in hindsight.

5.4.4 Notes

The construction of the non-Lebesgue measurable subset of Example 5.4.3 is due to Vitali.

Exercises

5.4.1 Using the definition of the Lebesgue measure show that the measure of a singleton is zero.

5.4.2 Let $A \subseteq \mathbb{R}$ be Lebesgue measurable and for $\rho \in \mathbb{R}_{>0}$ define

$$\rho A = \{\rho x \mid x \in A\}.$$

Show that $\lambda(\rho A) = \rho\lambda(A)$.

5.4.3 For the following subsets of $\mathbb{R}$, verify that they are Borel subsets (and therefore measurable sets), and determine their Lebesgue measure:

(a) the bounded, open interval (a, b);

(b) the bounded, open-closed interval $(a, b]$;

(c) the bounded, closed-open interval $[a, b)$;

(d) the singleton $\{x\}$ for any $x \in \mathbb{R}$;

(e) the unbounded closed interval $[a, \infty)$;

(f) the unbounded open interval (a, ∞).

5.4.4 Show that the set $\mathbb{Q}$ of rational numbers is a Borel set.

5.4.5 Show that G_δ's and F_σ's are Borel sets.

5.4.6 Show that for $x \in \mathbb{R}$, the point mass $\delta_x \colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ is regular.

5.4.7 Show that the counting measure $\mu \colon \mathscr{B}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ is not regular.

Section 5.5

Lebesgue measure on $\mathbb{R}^{\mathrm{n}}$

Although we will make most use of the Lebesgue measure on $\mathbb{R}$, we shall certainly have occasion to refer to the Lebesgue measure in higher dimensions, and so in this section we present this. The discussion here mirrors, for the most part, that in Section 5.4, so we will on occasion be a little sparse in our discussion.

Do I need to read this section? The material in this section can be bypassed until it is needed. •

5.5.1 The Lebesgue outer measure and the Lebesgue measure on $\mathbb{R}^{\mathrm{n}}$

As with the Lebesgue measure on $\mathbb{R}$, we construct the Lebesgue measure on $\mathbb{R}^n$ by first defining an outer measure. It is convenient to first define the volume of a rectangle. If $R = I_1 \times \cdots \times I_n$ is a rectangle in $\mathbb{R}^n$ we define its ***volume*** to be

$$\nu(R) = \begin{cases} \prod_{j=1}^{n} \ell(I_j), & \ell(I_j) < \infty,\ j \in \{1, \dots, n\}, \\ \infty, & \text{otherwise.} \end{cases}$$

With this notation we have the following definition.

5.5.1 Definition (Lebesgue outer measure on $\mathbb{R}^{\mathrm{n}}$) The ***Lebesgue outer measure*** on $\mathbb{R}^n$ is defined by

$$\lambda_n^*(S) = \inf\Big\{ \sum_{j=1}^{\infty} \nu(R_j) \ \Big|\ S \subseteq \bigcup_{j \in \mathbb{Z}_{>0}} R_j,\ R_j \text{ an open bounded rectangle, } j \in \mathbb{Z}_{>0} \Big\}. \quad \bullet$$

The Lebesgue outer measure on $\mathbb{R}^n$ has the same sort of naturality property with respect to volumes of rectangles that the Lebesgue outer measure on $\mathbb{R}$ has with respect to lengths of intervals.

5.5.2 Theorem (Lebesgue outer measure is an outer measure) *The Lebesgue outer measure on $\mathbb{R}^{\mathrm{n}}$ is an outer measure. Furthermore, if $\mathrm{R} = \mathrm{I}_1 \times \cdots \times \mathrm{I}_{\mathrm{n}}$ is a rectangle then $\lambda^*_{\mathrm{n}}(\Lambda) = \nu(\Lambda)$.*

Proof First we show that $\lambda_n^*(\emptyset) = 0$. Indeed, let $(\epsilon_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence converging to zero in $\mathbb{R}_{>0}$ and note that $\emptyset \subseteq (-\epsilon_j, \epsilon_j)^n$, $j \in \mathbb{Z}_{>0}$. Since $\lim_{j\to\infty} |\epsilon_j + \epsilon_j|^n = 0$, our assertion follows.

Next we show monotonicity of λ_n^*. This is clear since if $A \subseteq B \subseteq \mathbb{R}^n$ and if a collection of bounded open rectangles $(R_j)_{j \in \mathbb{Z}_{>0}}$ covers B, then the same collection of intervals covers A.

For countable-subadditivity of λ_n^*, let $(A_j)_{j \in \mathbb{Z}_{>0}}$ be a collection of subsets of $\mathbb{R}^n$. If $\sum_{j=1}^{\infty} \lambda_n^*(A_j) = \infty$ then countable-subadditivity follows trivially in this case, so we may as well suppose that $\sum_{j=1}^{\infty} \lambda_n^*(A_j) < \infty$. For $j \in \mathbb{Z}_{>0}$ and $\epsilon \in \mathbb{R}_{>0}$ let $(R_{j,k})_{k \in \mathbb{Z}_{>0}}$ be a collection of bounded open rectangles covering A_j and for which

$$\sum_{k=1}^{\infty} \nu(R_{j,k}) < \lambda_n^*(A_j) + \frac{\epsilon}{2^j}.$$

By Proposition **??**, $\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ is countable. Therefore, we may arrange the rectangles $(R_{j,k})_{j,k \in \mathbb{Z}_{>0}}$ into a single sequence $(R_l)_{l \in \mathbb{Z}_{>0}}$ so that

1. $\cup_{j \in \mathbb{Z}_{>0}} A_j \subseteq \cup_{l \in \mathbb{Z}_{>0}} R_l$ and
2. $\displaystyle\sum_{l=1}^{\infty} \nu(R_l) < \sum_{l=1}^{\infty} \Big(\lambda_n^*(A_l) + \frac{\epsilon}{2^l}\Big) = \sum_{l=1}^{\infty} \lambda_n^*(A_j) + \epsilon.$

This shows that

$$\lambda_n^*\Big(\bigcup_{j \in \mathbb{Z}_{>0}} A_j\Big) \leq \sum_{j=1}^{\infty} \lambda_n^*(A_j),$$

giving countable-subadditivity of λ_n^*.

We finally show that $\lambda^*(R) = \nu(R)$ for any rectangle R. We first take R to be compact. Let R_ϵ be an open rectangle containing R and for which $\nu(R_\epsilon) = \nu(R) + \epsilon$. Then

$$R \subseteq R_\epsilon \cup \left(\cup_{j=2}^{\infty} R_j\right),$$

where $R_j = \emptyset$, $j \geq 2$. Thus we have $\lambda_n^*(R) < \nu(R) + \epsilon$, and since this holds for every $\epsilon \in \mathbb{R}_{>0}$ it follows that $\nu(R) \leq \lambda_n^*(R)$. Now suppose that $(R_j)_{j \in \mathbb{Z}_{>0}}$ is a family of bounded open rectangles for which $R \subseteq \cup_{j \in \mathbb{Z}_{>0}} R_j$. Since R is compact, there is a finite subset of these rectangles, let us abuse notation slightly and denote them by $(R_1, \dots, R_k)$, such that $R \subseteq \cup_{j=1}^{k} R_j$. Now let $\boldsymbol{P}$ be a partition of R such that each of the subrectangles of $\boldsymbol{P}$ is contained in one of the rectangles $R_1, \dots, R_n$. This is possible since there are only finitely many of the rectangles $R_1, \dots, R_n$. By definition of the volume of a rectangle we have

$$\nu(R) = \sum_{R' \in \boldsymbol{P}} \nu(R') \leq \sum_{j=1}^{k} \nu(R_j) = \sum_{j=1}^{k} \lambda_n^*(R_j).$$

This gives $\nu(R) = \lambda_n^*(R)$, as desired.

Now let R be a bounded rectangle. Since $R \subseteq \mathrm{cl}(R)$ we have $\lambda_n^*(R) \leq \nu(\mathrm{cl}(R)) = \nu(R)$ using monotonicity of λ_n^*. If $\epsilon \in \mathbb{R}_{>0}$ we may find a compact rectangle $R_\epsilon \subseteq R$ for which $\nu(R) \leq \nu(R_\epsilon) + \epsilon$. Since $\lambda_n^*(R_\epsilon) \leq \lambda_n^*(R)$ by monotonicity, it follows that

$$\lambda^*(R) \geq \lambda^*(R_\epsilon) = \nu(R_\epsilon) \geq \nu(R) - \epsilon.$$

Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary $\lambda_n^*(R) \geq \nu(R)$, showing that $\lambda_n^*(R) = \nu(R)$, as desired.

Finally, if R is unbounded then for any $M \in \mathbb{R}_{>0}$ we may find a compact rectangle $R' \subseteq R$ for which $\lambda_n^*(R') > M$. Since $\lambda_n^*(R) \geq \lambda_n^*(R')$ by monotonicity this means that $\lambda_n^*(R) = \infty$. ■

As with the Lebesgue outer measure on $\mathbb{R}$, there are subsets of $\mathbb{R}^n$ that are not Lebesgue measurable.

5.5.3 Example (A set that is not λ_n^*-measurable) Let $A \subseteq (0,1)$ be the subset of $\mathbb{R}$ constructed in Example 5.4.3 that is not λ^*-measurable. Then define $A_n = A \times (0,1) \times \cdots \times (0,1) \subseteq \mathbb{R}^n$. Then recall from Example 5.4.3 that $(0,1)$ is a countable union of translates of A. Thus $(0,1)^n$ is a countable union of translates of A_n. Since λ_n^* is translation-invariant as we shall show in Theorem 5.5.22, it follows that, if A_n is λ_n^*-measurable, then we have the same dichotomy for A_n as we had for A:

1. if $\lambda_n^*(A_n) = 0$ then $\lambda_n^*((0,1)^n) = 0$;
2. if $\lambda_n^*(A_n) \in \mathbb{R}_{>0}$ then $\lambda_n^*((0,1)^n) = \infty$.

Since both of these conclusions are false, it must be the case that A_n is not λ_n^*-measurable. •

This then leads to the following definition.

5.5.4 Definition (Lebesgue measurable subsets of $\mathbb{R}^n$, Lebesgue measure on $\mathbb{R}^n$) Let λ_n^* be the Lebesgue outer measure on $\mathbb{R}^n$ and denote by $\mathscr{L}(\mathbb{R}^n)$ the set of λ_n^*-measurable subsets of $\mathbb{R}^n$. The sets in $\mathscr{L}(\mathbb{R}^n)$ are called ***Lebesgue measurable***, or merely ***measurable***, and the complete measure $\lambda_n \colon \mathscr{L}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ induced by λ_n^* is called the ***Lebesgue measure*** on $\mathbb{R}^n$. •

As with the Lebesgue measure on $\mathbb{R}$, the Lebesgue measure on $\mathbb{R}^n$ can be restricted to measurable sets.

5.5.5 Proposition (Restriction of Lebesgue measure to measurable sets) *Let* $\mathrm{A} \in \mathscr{L}(\mathbb{R}^\mathrm{n})$ *and denote*

(i) $\mathscr{L}(\mathrm{A}) = \{\mathrm{B} \cap \mathrm{A} \mid \mathrm{B} \in \mathscr{L}(\mathbb{R}^\mathrm{n})\}$ *and*

(ii) $\lambda_\mathrm{A} \colon \mathscr{L}(\mathrm{A}) \to \overline{\mathbb{R}}_{\geq 0}$ *given by* $\lambda_\mathrm{A}(\mathrm{B} \cap \mathrm{A}) = \lambda(\mathrm{B} \cap \mathrm{A})$.

Then $(\mathrm{A}, \mathscr{L}(\mathrm{A}), \lambda_\mathrm{A})$ *is a complete measure space.*

Proof This follows from Propositions 5.2.6, 5.3.18, and 5.3.23. ■

5.5.2 Borel sets in $\mathbb{R}^n$ as examples of Lebesgue measurable sets

Next we turn to the Borel sets in $\mathbb{R}^n$ which provide a large and somewhat comprehensible collection of Lebesgue measurable sets. We denote by $\mathscr{O}(\mathbb{R}^n)$ the open subsets of $\mathbb{R}^n$.

5.5.6 Definition (Borel subsets of $\mathbb{R}^n$) The collection of ***Borel sets*** in $\mathbb{R}^n$ is the σ-algebra generated by $\mathscr{O}(\mathbb{R}^n)$ (see Proposition 5.2.7). We denote by $\mathscr{B}(\mathbb{R}^n)$ the Borel sets in $\mathbb{R}^n$. If $A \in \mathscr{B}(\mathbb{R}^n)$ then we denote

$$\mathscr{B}(A) = \{A \cap B \mid B \in \mathscr{B}(\mathbb{R}^n)\}$$ •

While it is not so easy to come up with a satisfactory description of *all* Borel sets, it is the case that we will only encounter non-Borel sets as examples of things that are peculiar. Thus one can frequently get away with only thinking of Borel sets when one thinks about Lebesgue measurable sets. We shall be a little more precise about just what this means later.

For the moment, let us give a few examples of Borel sets. The following result gives us a ready made and very large class of Borel sets. In the following result we make the natural identification of $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ with $\mathbb{R}^{n_1+n_2}$.

5.5.7 Proposition (Products of Borel sets) *Let $\sigma(\mathscr{B}(\mathbb{R}^{\mathrm{n}_1}) \times \mathscr{B}(\mathbb{R}^{\mathrm{n}_2}))$ denote the σ-algebra on $\mathbb{R}^{\mathrm{n}_1} \times \mathbb{R}^{\mathrm{n}_2}$ generated by subsets of the form $\mathrm{B}_1 \times \mathrm{B}_2$, where $\mathrm{B}_1 \in \mathscr{B}(\mathbb{R}^{\mathrm{n}_1})$ and $\mathscr{B}(\mathbb{R}^{\mathrm{n}_2})$. Then $\mathscr{B}(\mathbb{R}^{\mathrm{n}_1+\mathrm{n}_2}) = \sigma(\mathscr{B}(\mathbb{R}^{\mathrm{n}_1}) \times \mathscr{B}(\mathbb{R}^{\mathrm{n}_2}))$.*

Proof By ***missing stuff*** it follows that the open sets in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ are countable unions of sets of the form $U_1 \times U_2$ where $U_1 \subseteq \mathbb{R}^{n_1}$ and $U_2 \subseteq \mathbb{R}^{n_2}$ are open. By Exercise 5.2.4 it follows that $\mathscr{B}(\mathbb{R}^{n_1+n_2})$ is generated by subsets from the σ-algebra $\sigma(\mathscr{B}(\mathbb{R}^{n_1}) \times \mathscr{B}(\mathbb{R}^{n_2}))$. Thus $\mathscr{B}(\mathbb{R}^{n_1+n_2}) \subseteq \sigma(\mathscr{B}(\mathbb{R}^{n_1}) \times \mathscr{B}(\mathbb{R}^{n_2}))$.

For the converse inclusion, note that the projections $\mathrm{pr}_1 \colon \mathbb{R}^{n_1+n_2} \to \mathbb{R}^{n_1}$ and $\mathrm{pr}_2 \colon \mathbb{R}^{n_1+n_2} \to \mathbb{R}^{n_2}$ are continuous. From this one can easily show (and this will be shown in Example 5.5.10–3) that $\pi_1^{-1}(B_1), \mathrm{pr}_2^{-1}(B_2) \in \mathscr{B}(\mathbb{R}^{n_1+n_2})$ for $B_1 \in \mathscr{B}(\mathbb{R}^{n_1})$ and $B_2 \in \mathscr{B}(\mathbb{R}^{n_2})$. Therefore,

$$B_1 \cap B_2 = \pi_1^{-1}(B_1) \cap \mathrm{pr}_2^{-1}(B_2) \in \mathscr{B}(\mathbb{R}^{n_1+n_2})$$

for $B_1 \in \mathscr{B}(\mathbb{R}^{n_1})$ and $B_2 \in \mathscr{B}(\mathbb{R}^{n_2})$. Thus $\sigma(\mathscr{B}(\mathbb{R}^{n_1}) \times \mathscr{B}(\mathbb{R}^{n_2})) \subseteq \mathscr{B}(\mathbb{R}^{n_1+n_2})$ since $\sigma(\mathscr{B}(\mathbb{R}^{n_1}) \times \mathscr{B}(\mathbb{R}^{n_2}))$ is the smallest σ-algebra containing products of Borel sets. ■

5.5.8 Remark ($\sigma(\mathscr{L}(\mathbb{R}^{\mathrm{n}_1}) \times \mathscr{L}(\mathbb{R}^{\mathrm{n}_2})) \neq \mathscr{L}(\mathbb{R}^{\mathrm{n}_1+\mathrm{n}_2})$) The reader will notice that the result analogous to the preceding one, but for Lebesgue measurable sets was not stated. This is because it is actually not true, as will be seen ***missing stuff***. This is an instance that illustrates that the mantra "What seems like it should be true is true" should always be verified explicitly. •

The following alternative characterisations of Borel sets are sometimes useful.

5.5.9 Proposition (Alternative characterisation of Borel sets) *$\mathscr{B}(\mathbb{R}^{\mathrm{n}})$ is equal to the following collections of sets:*

(i) the σ-algebra $\mathscr{B}_1$ generated by the closed subsets;

(ii) the σ-algebra $\mathscr{B}_2$ generated by rectangles of the form $(-\infty, \mathrm{b}_1] \times \cdots \times (-\infty, \mathrm{b}_\mathrm{n}]$, $\mathrm{b}_1, \dots, \mathrm{b}_\mathrm{n} \in \mathbb{R}$;

(iii) the σ-algebra $\mathscr{B}_3$ generated by intervals of the form $(\mathrm{a}_1, \mathrm{b}_1] \times \cdots \times (\mathrm{a}_\mathrm{n}, \mathrm{b}_\mathrm{n}]$, $\mathrm{a}_\mathrm{j}, \mathrm{b}_\mathrm{j} \in \mathbb{R}$, $\mathrm{a}_\mathrm{j} < \mathrm{b}_\mathrm{n} \in \mathbb{R}$, $\mathrm{j} \in \{1, \dots, \mathrm{n}\}$.

Proof First note that $\mathscr{B}(\mathbb{R}^n)$ contains the σ-algebra $\mathscr{B}_1$ generated by all closed sets, since the complements of all open sets, i.e., all closed sets, are contained in $\mathscr{B}(\mathbb{R}^n)$. Note that the sets of the form $(-\infty, b_1] \times \cdots \times (-\infty, b_n]$ are closed, so the σ-algebra $\mathscr{B}_2$ generated by these subsets is contained in $\mathscr{B}_1$. Since $(a_j, b_j] = (-\infty, b_j] \cap (\mathbb{R} \setminus (-\infty, a_j])$, $j \in \{1, \dots, n\}$, it follows that the σ-algebra $\mathscr{B}_3$ is contained in $\mathscr{B}_2$. Finally, note that

$$(a_j, b_j) = \cup_{k=1}^{\infty} (a_j, b_k - \tfrac{1}{k}], \qquad j \in \{1, \dots, n\}.$$

Thus, by ***missing stuff***, each open subset of $\mathbb{R}^n$ is a countable union of sets, each of which is a countable intersection of generators of sets of $\mathscr{B}_3$. Thus $\mathscr{B}(\mathbb{R}^n) \subseteq \mathscr{B}_3$. Putting this all together gives

$$\mathscr{B}(\mathbb{R}^n) \subseteq \mathscr{B}_3 \subseteq \mathscr{B}_2 \subseteq \mathscr{B}_1 \subseteq \mathscr{B}(\mathbb{R}^n).$$

Thus we must conclude that $\mathscr{B}_1 = \mathscr{B}_2 = \mathscr{B}_3 = \mathscr{B}(\mathbb{R}^n)$. ■

We can now give some examples of Borel sets in $\mathbb{R}^n$.

5.5.10 Examples (Borel sets)

1. We claim that if $B_1, \dots, B_n \in \mathscr{B}(\mathbb{R})$ then $B_1 \times \cdots \times B_n \in \mathscr{B}(\mathbb{R}^n)$; this follows by a simple induction from Proposition 5.5.7. This provides us with a large collection of Borel sets, provided we have Borel sets in $\mathbb{R}$.
2. As for Borel sets in $\mathbb{R}$, a set that is a countable intersection of open sets is called a $\mathbf{G}_\delta$ and a set that is a countable union of closed sets is called an $\mathbf{F}_\sigma$.
3. If $B \in \mathscr{B}(\mathbb{R}^n)$ and if $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous, then $f^{-1}(B) \in \mathscr{B}(\mathbb{R}^n)$. To see, by Proposition 5.5.9 it suffices to show that
$$f^{-1}((-\infty, b_1] \times \cdots \times (-\infty, b_n])$$
is closed. If $f(x) = (f_1(x, \dots, f_m(x))$ then
$$f^{-1}((-\infty, b_1] \times \cdots \times (-\infty, b_n]) = f_1^{-1}((-\infty, b_1]) \cap \cdots \cap f_n^{-1}((-\infty, b_n]).$$
Since each of the functions $f_1, \dots, f_n$ are continuous it follows from Corollary **??** that $f_j^{-1}((-\infty, b_n])$ is closed for each $j \in \{1, \dots, n\}$. Thus
$$f^{-1}((-\infty, b_1] \times \cdots \times (-\infty, b_n])$$
is closed, being a finite intersection of closed sets. This gives the desired conclusion.

This again gives us a wealth of Borel sets. •

Now that we understand a little of the character of Borel sets, let us provide their relationship with the Lebesgue measurable sets. As with the relationship of $\mathscr{B}(\mathbb{R})$ with $\mathscr{L}(\mathbb{R})$, the correspondence between Borel and Lebesgue measurable sets in $\mathbb{R}^n$ has its nice points and its somewhat deficient aspects.

5.5.11 Theorem (Borel sets are Lebesgue measurable) $\mathscr{B}(\mathbb{R}^n) \subseteq \mathscr{L}(\mathbb{R}^n)$.

Proof The theorem will follow from Proposition 5.5.9 if we can show that any set of the form $(-\infty, b_1] \times \cdots \times (-\infty, b_n]$ is Lebesgue measurable. Let A be such a set and note that since
$$\lambda_n^*(S) \leq \lambda_n^*(S \cap A) + \lambda_n^*(S \cap (\mathbb{R}^n \setminus A))$$
we need only show the opposite inequality to show that A is Lebesgue measurable. If $\lambda_n^*(S) = \infty$ this is clearly true, so we may as well suppose that $\lambda_n^*(S) < \infty$. Let $(R_j)_{j \in \mathbb{Z}_{>0}}$ be bounded open rectangles that cover S and be such that
$$\sum_{j=1}^{\infty} \nu(R_j) < \lambda_n^*(S) + \epsilon.$$
For $j \in \mathbb{Z}_{>0}$ choose bounded open rectangles D_j and E_j, possibly empty, for which
$$\begin{aligned}&R_j \cap A \subseteq D_j,\\ &R_j \cap (\mathbb{R}^n \setminus A) \subseteq E_j,\\ &\nu(D_j) + \nu(E_j) \leq \nu(R_j) + \frac{\epsilon}{2^j}.\end{aligned}$$

Note that the bounded open rectangles $(D_j)_{j\in\mathbb{Z}_{>0}}$ cover $S\cap A$ and that the bounded open rectangles $(E_j)_{j\in\mathbb{Z}_{>0}}$ cover $\mathbb{R}^n\setminus A$ so that

$$\lambda_n^*(S\cap A)\le\sum_{j=1}^\infty \nu(D_j),\quad \lambda_n^*(S\cap(\mathbb{R}^n\setminus A))\le\sum_{j=1}^\infty \nu(E_j).$$

From this we have

$$\lambda_n^*(S\cap A)+\lambda_n^*(S\cap(\mathbb{R}^n\setminus A))\le\sum_{j=1}^\infty \nu(R_j)+\epsilon<\lambda_n^*(S)+2\epsilon,$$

using the fact that $\sum_{j=1}^\infty \frac{1}{2^j}=1$ by Example 2.4.2–**??**. Since ϵ can be taken arbitrarily small, the inequality

$$\lambda_n^*(S)\ge\lambda_n^*(S\cap A)+\lambda_n^*(S\cap(\mathbb{R}\setminus A))$$

follows, and so too does the result. ■

While the preceding result is useful in that it tells us that the large class of (sort of) easily understood Borel sets are Lebesgue measurable, the following result says that much more is true. Namely, up to sets of measure zero, all Lebesgue measurable sets are Borel sets.

5.5.12 Theorem (Lebesgue measurable sets are the completion of the Borel sets) $(\mathbb{R}^n,\mathscr{L}(\mathbb{R}^n),\lambda_n)$ *is the completion of* $(\mathbb{R}^n,\mathscr{B}(\mathbb{R}^n),\lambda_n|\mathscr{B}(\mathbb{R}^n))$.

Proof First, given $A\in\mathscr{L}(\mathbb{R}^n)$, we find $L,U\in\mathscr{B}(\mathbb{R}^n)$ such that $L\subseteq A\subseteq U$ and such that $\lambda_n(U\setminus L)=0$. We first suppose that $\lambda_n(A)<\infty$. Using Theorem 5.5.18 below, let $(U_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of open sets containing A and for which $\lambda_n(U_j)\le\lambda_n(A)+\frac{1}{j}$ and let $(L_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of compact subsets of A for which $\lambda_n(L_j)\ge\lambda_n(A)-\frac{1}{j}$. If we take $L=\cup_{j\in\mathbb{Z}_{>0}}L_j$ and $U=\cap_{j\in\mathbb{Z}_{>0}}U_j$ then we have $L\subseteq A\subseteq U$. We also have

$$\lambda_n(U\setminus L)\le\lambda_n(U_j\setminus L_j)=\lambda_n(U_j\setminus A)+\lambda_n(A\setminus L_j)\le\tfrac{1}{2j}.$$

Since this holds for every $j\in\mathbb{Z}_{>0}$, this gives our claim when A has finite measure, since L and U are Borel sets. If $\lambda_n(A)=\infty$ then we can write $A=\cup_{j\in\mathbb{Z}_{>0}}A_j$ with $A_j=(-j,j)^n\cap A$. For each $j\in\mathbb{Z}_{>0}$ we may find $L_j,U_j\in\mathscr{B}(\mathbb{R})$ such that $L_j\subseteq A_j\subseteq U_j$ and $\lambda_n(U_j\setminus L_j$. Taking $L=\cup_{j\in\mathbb{Z}_{>0}}L_j$ and $U=\cup_{j\in\mathbb{Z}_{>0}}$ gives $L\subseteq A\subseteq U$ and $\lambda_n(U\setminus L)=0$.

The above shows that $\mathscr{L}(\mathbb{R}^n)\subseteq\mathscr{B}_{\lambda_n}(\mathbb{R}^n)$. Now let $B\in\mathscr{B}_{\lambda_n}(\mathbb{R})$ and take Borel sets L and U for which $L\subseteq B\subseteq U$ and $\lambda_n(U\setminus L)=0$. Note that $(B\setminus L)\subseteq(U\setminus L)$. Note also that since $U\setminus L\in\mathscr{B}(\mathbb{R})$ we have $U\setminus L\in\mathscr{L}(\mathbb{R})$ and $\lambda_n(U\setminus L)=0$. By completeness of the Lebesgue measure this implies that $B\setminus L\in\mathscr{L}(\mathbb{R})$. Since $B=(B\setminus L)\cup L$ this implies that $B\in\mathscr{L}(\mathbb{R})$. ■

The theorem has the following corollary which explicitly indicates what it means to approximate a Lebesgue measurable set with a Borel set.

5.5.13 Corollary (Borel approximations to Lebesgue measurable sets) *If* $A\in\mathscr{L}(\mathbb{R}^n)$ *then there exists a Borel set* B *and a set* Z *of measure zero such that* $A=B\cup Z$.

Proof This follows directly from Theorem 5.5.12 and the definition of the completion. ■

As is the case for $\mathscr{B}(\mathbb{R})$ and $\mathscr{L}(\mathbb{R})$, there are many more sets in $\mathscr{L}(\mathbb{R}^n)$ than there are in $\mathscr{B}(\mathbb{R}^n)$, the preceding corollary notwithstanding.

5.5.14 Proposition (The cardinalities of Borel and Lebesgue measurable sets) *We have* $\mathrm{card}(\mathscr{B}(\mathbb{R}^n)) = \mathrm{card}(\mathbb{R})$ *and* $\mathrm{card}(\mathscr{L}(\mathbb{R}^n)) = \mathrm{card}(\mathbf{2}^{\mathbb{R}})$.

Proof Since $\{x\} \in \mathscr{B}(\mathbb{R}^n)$ for every $x \in \mathbb{R}^n$ we obviously have $\mathrm{card}(\mathscr{B}(\mathbb{R}^n)) \geq \mathrm{card}(\mathbb{R}^n) = \mathrm{card}(\mathbb{R})$, the last equality holding by virtue of Theorem **??**. For the opposite inequality, note that Proposition **??** it holds that every open set is a union of open balls with rational radius and whose centres have rational coordinates in $\mathbb{R}^n$. There are countable many such balls by Proposition **??**. Let $\mathscr{S}$ be the set of such balls and note that, adopting the notation of Theorem 5.2.14, $\mathscr{S}_1$ therefore includes the open subsets of $\mathbb{R}^n$. Thus $\mathscr{B}(\mathbb{R}^n)$ is the σ-algebra generated by $\mathscr{S}$ and so, by Theorem 5.2.14, $\mathrm{card}(\mathscr{B}(\mathbb{R}^n)) \leq \aleph_0^{\aleph_0}$. Since

$$2^{\aleph_0} \leq \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0},$$

using the fact that $2 \leq \aleph_0 \leq 2^{\aleph_0}$ by Example **??–??** and Exercise **??**, it follows that $\mathrm{card}(\mathscr{B}(\mathbb{R}^n)) \leq \mathrm{card}(\mathbb{R})$, as desired.

Next, we obviously have

$$\mathrm{card}(\mathscr{L}(\mathbb{R}^n)) \leq \mathrm{card}(\mathbf{2}^{\mathbb{R}^n}) = \mathrm{card}(\mathbf{2}^{\mathbb{R}}),$$

using the fact that $\mathrm{card}(\mathbb{R}^n) = \mathrm{card}(\mathbb{R})$ by Theorem **??**. For the opposite inequality, we note that the Cantor set $C \subseteq [0, 1]$ has Lebesgue measure zero and has the cardinality of $[0, 1]$, and thus the cardinality of $\mathbb{R}$. Thus the set $C_n = C \times \mathbb{R}^{n-1} \subseteq \mathbb{R}^n$ also has measure zero (why?), and satisfies

$$\mathrm{card}(C_n) = \mathrm{card}(C) \cdot \mathrm{card}(\mathbb{R}^n) = \mathrm{card}(\mathbb{R})^n = \mathrm{card}(\mathbb{R}),$$

using Theorem **??**. Since $\mathscr{L}(\mathbb{R}^n)$ is complete it follows that every subset of C_n is Lebesgue measurable, and so

$$\mathrm{card}(\mathscr{L}(\mathbb{R}^n)) \geq \mathrm{card}(\mathbf{2}^{C_n}) = \mathrm{card}(\mathbf{2}^{\mathbb{R}}).$$

Thus $\mathrm{card}(\mathscr{L}(\mathbb{R}^n)) = \mathrm{card}(\mathbf{2}^{\mathbb{R}})$, as desired. ■

Using the fact that this is possible when $n = 1$, it is possible to construct a Lebesgue measurable subset of $\mathbb{R}^n$ that is not Borel.

5.5.15 Example (A non-Borel Lebesgue measurable set) Let $B \subseteq [0, 1]$ be the subset defined in Example 5.4.14, recalling that B is Lebesgue measurable but not Borel. We claim that $B_n = B \times \mathbb{R}^{n-1} \subseteq \mathbb{R}^n$ is Lebesgue measurable but not Borel. It is Lebesgue measurable since B is a subset of the Cantor set C which has zero measure, and so $B_n \subseteq C \times \mathbb{R}^{n-1}$ with $C \times \mathbb{R}^{n-1}$ having zero measure. Completeness of $\mathscr{L}(\mathbb{R}^n)$ ensures that B_n is Lebesgue measurable. However, B_n cannot be a Borel set. Indeed, let $i_1 \colon \mathbb{R} \to \mathbb{R}^n$ be the continuous map $i_1(x) = (x, 0, \dots, 0)$. Then one can easily see that $B = i_1^{-1}(B_n)$. Were B_n a Borel set, this would imply that B is a Borel set by Example 5.5.10–3. •

5.5.3 Further properties of the Lebesgue measure on $\mathbb{R}^n$

In this section we shall establish some important properties of the Lebesgue measure. These are intended to show the extent to which the Lebesgue measure is a natural and well-behaved construction.

We begin with an important attribute of measures in general.

5.5.16 Definition (Regular measure on $\mathbb{R}^{\mathrm{n}}$) Let $\mathscr{A}$ be a σ-algebra on $\mathbb{R}^n$ that contains the Borel σ-algebra $\mathscr{B}(\mathbb{R}^n)$. A measure $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ is ***regular*** if

(i) $\mu(K) < \infty$ for each compact subset $K \subseteq \mathbb{R}^n$,

(ii) if $A \in \mathscr{A}$ then $\mu(A) = \inf\{\mu(U) \mid U \text{ open and } A \subseteq U\}$, and

(iii) if $U \subseteq \mathbb{R}^n$ is open then $\mu(U) = \sup\{\mu(K) \mid K \text{ open and } K \subseteq U\}$. •

Let us give some simple examples to illustrate what regular means.

5.5.17 Examples (Regular and irregular measures)

1. If $x \in \mathbb{R}^n$, the point mass measure $\delta_x\colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$\delta(B) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B \end{cases}$$

is regular, as may be readily verified; see Exercise 5.5.2.

2. One can check that the counting measure $\mu\colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ defined by

$$\mu(B) = \begin{cases} \operatorname{card}(B), & \operatorname{card}(B) < \infty, \\ \infty, & \text{otherwise} \end{cases}$$

is not regular; see Exercise 5.5.3. •

5.5.18 Theorem (Regularity of the Lebesgue measure) *The Lebesgue measure $\lambda_{\mathrm{n}}\colon \mathscr{L}(\mathbb{R}^{\mathrm{n}}) \to \overline{\mathbb{R}}_{\geq 0}$ is σ-finite and regular. Moreover, for $\mathrm{A} \in \mathscr{L}(\mathbb{R}^{\mathrm{n}})$ we have $\lambda_{\mathrm{n}}(\mathrm{A}) = \sup\{\lambda + \mathrm{n}(\mathrm{K}) \mid K \text{ compact and } K \subseteq A\}$.*

Proof To see that λ_n is σ-finite note that $\mathbb{R}^n = \cup_{k\in\mathbb{Z}_{>0}}[-k,k]^n$ with $\lambda_n([-k,k]^n) < \infty$.

Next we show that if $A \in \mathscr{L}(\mathbb{R}^n)$ then

$$\lambda_n(A) = \inf\{\lambda_n(U) \mid U \text{ open and } A \subseteq U\}.$$

Assume that $\lambda_n(A) < \infty$ since the result is obvious otherwise. Let $\epsilon \in \mathbb{R}_{>0}$ and let $(R_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of bounded open rectangles for which $A \subseteq \cup_{j\in\mathbb{Z}_{>0}} R_j$ and for which

$$\sum_{j=1}^{\infty} \nu(R_j) = \lambda_n(A) + \epsilon.$$

Now let $U = \cup_{j\in\mathbb{Z}_{>0}} R_j$, noting that U is open and that $A \subseteq U$. By Proposition 5.3.10(iii) and the fact that the measure of a rectangle is its we have

$$\lambda_n(U) \leq \sum_{j=1}^{\infty} \nu(R_j) = \lambda_n(A) + \epsilon.$$

Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary this shows that

$$\lambda_n(A) \geq \inf\{\lambda_n(U) \mid U \text{ open and } A \subseteq U\}.$$

Since the other inequality is obvious by the basic properties of a measure, this part of the result follows.

Note that to show that λ_n is regular it suffices to prove the final assertion of the theorem since open sets are Lebesgue measurable; thus we prove the final assertion of the theorem. First suppose that $A \in \mathscr{L}(\mathbb{R}^n)$ is bounded. Then let $\tilde{K}$ be a compact set containing A. For $\epsilon \in \mathbb{R}_{>0}$ choose U open and containing $\tilde{K} \setminus A$ and for which $\lambda_n(U) \leq \lambda_n(\tilde{K} \setminus A) + \epsilon$, this being possible from by the first part of the proof. Note that $K = \tilde{K} \setminus U$ is then a compact set contained in A and that the basic properties of measure then give

$$\lambda_n(U) \leq \lambda_n(\tilde{K} \setminus A) + \epsilon \ \text{ and } \ \lambda_n(\tilde{K}) \leq \lambda_n(K) - \lambda_n(A) \quad \Longrightarrow \quad \lambda_n(K) > \lambda_n(A) - \epsilon.$$

Since ϵ can be made as small as desired, this gives the second part of the proposition when A is bounded. Define

$$A_j = (-j, j)^n \cap A,$$

and note that $(A_j)_{j \in \mathbb{Z}_{>0}}$ is an increasing sequence of sets and that $A = \cup_{j \in \mathbb{Z}_{>0}} A_j$. Therefore, by Proposition 5.3.10(iv), $\lambda_n(A) = \lim_{j \to \infty} \lambda_n(A_j)$. Then for any $M < \lambda_n(A)$ there exists $N \in \mathbb{Z}_{>0}$ such that $\lambda_n(A_N) > M$. We may now find a compact K such that $\lambda_n(K) > M$ by the fact that we have proved our assertion for bounded sets (as is A_N). Note that $K \subseteq A$ and that $M < \lambda_n(A)$ is arbitrary, and so the result follows. ■

The theorem has the following corollary.

5.5.19 Corollary (Approximation of Lebesgue measurable sets by open and compact sets) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R}^\mathrm{n})$ *satisfies* $\lambda_\mathrm{n}(\mathrm{A}) < \infty$ *and if* $\epsilon \in \mathbb{R}_{>0}$ *then there exists an open set* $\mathrm{U} \subseteq \mathbb{R}^\mathrm{n}$ *and a compact set* $\mathrm{K} \subseteq \mathbb{R}^\mathrm{n}$ *such that*

$$\lambda_\mathrm{n}(\mathrm{U} \setminus \mathrm{A}) < \epsilon, \quad \lambda_\mathrm{n}(\mathrm{A} \setminus \mathrm{K}) < \epsilon.$$

Next we show that the Lebesgue measure has the quite natural property of being translation-invariant. First we provide definitions for translation-invariant σ-algebras and measures.

5.5.20 Definition (Translation-invariant σ-algebra and measure on $\mathbb{R}^\mathrm{n}$) A σ-algebra $\mathscr{A} \subseteq \mathbf{2}^{\mathbb{R}^n}$ is ***translation-invariant*** if, for every $A \in \mathscr{A}$ and every $\boldsymbol{x} \in \mathbb{R}^n$,

$$\boldsymbol{x} + A \triangleq \{\boldsymbol{x} + \boldsymbol{y} \mid \ \boldsymbol{y} \in A\} \in \mathscr{A}.$$

A ***translation-invariant*** measure on a translation-invariant σ-algebra $\mathscr{A}$ is a map $\mu \colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ for which $\mu(\boldsymbol{x} + A) = \mu(A)$ for every $A \in \mathscr{A}$ and $\boldsymbol{x} \in \mathbb{R}^n$. •

The Borel and Lebesgue measurable sets are translation-invariant.

5.5.21 Proposition (Translation-invariance of Borel and Lebesgue measurable sets) *Both* $\mathscr{B}(\mathbb{R}^\mathrm{n})$ *and* $\mathscr{L}(\mathbb{R}^\mathrm{n})$ *are translation-invariant.*

Proof Let us denote

$$\mathscr{B}'(\mathbb{R}^n) = \{B \mid \ \boldsymbol{x} + B \in \mathscr{B}(\mathbb{R}^n) \text{ for every } \boldsymbol{x} \in \mathbb{R}^n\}.$$

We claim that $\mathscr{B}'(\mathbb{R}^n)$ is a σ-algebra containing the open subsets of $\mathbb{R}^n$. First of all, if $U \subseteq \mathbb{R}^n$ is open then $\boldsymbol{x} + U$ is open for every $\boldsymbol{x} \in \mathbb{R}^n$ (why?) and so $U \in \mathscr{B}'(\mathbb{R}^n)$. To see that $\mathscr{B}'(\mathbb{R}^n)$ is a σ-algebra, first note that $\mathbb{R}^n = \boldsymbol{x} + \mathbb{R}^n$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and so $\mathbb{R}^n \in \mathscr{B}'(\mathbb{R}^n)$. Next, let $B \in \mathscr{B}'(\mathbb{R}^n)$ and let $\boldsymbol{x} \in \mathbb{R}^n$. Then

$$\begin{aligned}\boldsymbol{x} + (\mathbb{R}^n \setminus B) &= \{\boldsymbol{x} + \boldsymbol{z} \mid \boldsymbol{z} \notin B\} = \{\boldsymbol{y} \mid \boldsymbol{y} - \boldsymbol{x} \notin B\} = \{\boldsymbol{y} \mid \boldsymbol{y} \neq \boldsymbol{x} + \boldsymbol{z},\ \boldsymbol{z} \in B\}\\ &= \{\boldsymbol{y} \mid \boldsymbol{y} \notin (\boldsymbol{x} + B)\} = \mathbb{R}^n \setminus (\boldsymbol{x} + B) \in \mathscr{B}(\mathbb{R}^n).\end{aligned}$$

Thus $\boldsymbol{x} + (\mathbb{R}^n \setminus B) \in \mathscr{B}(\mathbb{R}^n)$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and so $\mathbb{R}^n \setminus B \in \mathscr{B}'(\mathbb{R}^n)$. Finally, let $(B_j)_{j\mathbb{Z}_{>0}}$ be a countable collection of subsets from $\mathscr{B}'(\mathbb{R}^n)$. Then, for $\boldsymbol{x} \in \mathbb{R}^n$ we have

$$\boldsymbol{x} + \cup_{j\in\mathbb{Z}_{>0}} B_j = \cup_{j\in\mathbb{Z}_{>0}} (\boldsymbol{x} + B_j) \in \mathscr{B}(\mathbb{R}^n)$$

and so $\cup_{j\in\mathbb{Z}_{>0}} B_j \in \mathscr{B}'(\mathbb{R}^n)$. Thus $\mathscr{B}'(\mathbb{R}^n)$ is indeed a σ-algebra containing the open sets and so we conclude that $\mathscr{B}(\mathbb{R}^n) \subseteq \mathscr{B}'(\mathbb{R}^n)$ since $\mathscr{B}(\mathbb{R}^n)$ is the σ-algebra generated by the open sets. This shows that $\mathscr{B}(\mathbb{R}^n)$ is translation-invariant.

Next let us show that $\mathscr{L}(\mathbb{R}^n)$ is translation-invariant. To do this we first show that if $S \subseteq \mathbb{R}^n$ and if $\boldsymbol{x} \in \mathbb{R}^n$ then $\lambda_n^*(\boldsymbol{x} + S) = \lambda_n^*(S)$. Indeed,

$$\begin{aligned}\lambda_n^*(\boldsymbol{x} + S) &= \inf\Big\{\sum_{j=1}^{\infty} \nu(R_j) \Bigm| \boldsymbol{x} + S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} R_j\Big\}\\ &= \inf\Big\{\sum_{j=1}^{\infty} \nu(R_j') \Bigm| \boldsymbol{x} + S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} \nu(\boldsymbol{x} + R_j')\Big\}\\ &= \inf\Big\{\sum_{j=1}^{\infty} \nu(R_j') \Bigm| S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} R_j'\Big\} = \lambda_n^*(S),\end{aligned}$$

using the fact that for a rectangle R we have $\nu(R) = \nu(\boldsymbol{x} + R)$. Now let $A \in \mathscr{L}(\mathbb{R}^n)$ so that, for every subset $S \subseteq \mathbb{R}^n$,

$$\lambda_n^*(S) = \lambda_n^*(S \cap A) + \lambda_n^*(S \cap (\mathbb{R}^n \setminus A)).$$

Then, for $\boldsymbol{x} \in \mathbb{R}^n$ and $S \subseteq \mathbb{R}^n$,

$$\lambda_n^*(S \cap (\boldsymbol{x} + A)) = \lambda_n^*((\boldsymbol{x} + (-\boldsymbol{x} + S)) \cap (\boldsymbol{x} + A)) = \lambda_n^*((-\boldsymbol{x} + S) \cap A)$$

and, similarly,

$$\lambda_n^*(S \cap (\mathbb{R}^n \setminus (\boldsymbol{x} + A))) = \lambda_n^*((\boldsymbol{x} + (-\boldsymbol{x} + S)) \cap (\boldsymbol{x} + \mathbb{R}^n \setminus A)) = \lambda_n^*((-\boldsymbol{x} + S) \cap (\mathbb{R}^n \setminus A)).$$

Since $\lambda_n^*(-\boldsymbol{x} + S) = \lambda_n^*(S)$ this immediately gives

$$\lambda_n^*(S) = \lambda_n^*(S \cap (\boldsymbol{x} + A)) + \lambda_n^*(S \cap (\mathbb{R}^n \setminus (\boldsymbol{x} + A))),$$

showing that $\boldsymbol{x} + A \in \mathscr{L}(\mathbb{R}^n)$. ■

We may also show that the Lebesgue measure is translation-invariant, and is, moreover, in some sense unique.

5.5.22 Theorem (Translation invariance of the Lebesgue measure) *If* $\mu\colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ *is a nonzero translation-invariant measure for which* $\mu(B) < \infty$ *for every bounded* $B \in \mathscr{B}(\mathbb{R}^n)$*, then there exists* $c \in \mathbb{R}_{>0}$ *such that* $\mu(B) = c\lambda_n(B)$ *for every* $B \in \mathscr{B}(\mathbb{R}^n)$*. Moreover, the Lebesgue measure* $\lambda_n\colon \mathscr{L}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ *is translation-invariant.*

Proof That λ_n is translation-invariant follows from the proof of Proposition 5.4.22 where we showed that $\lambda_n^*(\boldsymbol{x} + S) = \lambda_n^*(S)$ for every $S \subseteq \mathbb{R}^n$ and $\boldsymbol{x} \in \mathbb{R}^n$. To show that λ_n is, up to a positive scalar, the only translation-invariant measure we first prove two lemmata.

1 Lemma *If* $U \subseteq \mathbb{R}^n$ *is a nonempty open set, then there exists a countable collection of disjoint rectangles* $(R_j)_{j\in\mathbb{Z}_{>0}}$ *of the form*

$$R_j = [a_{j,1}, b_{j,1}) \times \cdots \times [a_{j,n}, b_{j,n})$$

such that $U = \cup_{j\in\mathbb{Z}_{>0}} R_j$.

Proof For $k \in \mathbb{Z}_{\geq 0}$ define

$$\mathscr{C}_k = \{[j_1 2^{-k}, (j_1+1)2^{-k}) \times \cdots \times [j_n 2^{-k}, (j_n+1)2^{-k}) \mid j_1, \dots, j_n \in \mathbb{Z}\}.$$

Note that, for each $k \in \mathbb{Z}_{\geq 0}$, the sets from $\mathscr{C}_k$ form a countable partition of $\mathbb{R}^n$. Also note that for $k < l$, every cube in $\mathscr{C}_l$ is also a cube in $\mathscr{C}_k$. Now let $U \subseteq \mathbb{R}^n$ be open. Let $\mathscr{D}_0 = \emptyset$. Let

$$\begin{aligned}
\mathscr{D}_1 &= \{C \in \mathscr{C}_1 \mid C \subseteq U\},\\
\mathscr{D}_2 &= \{C \in \mathscr{C}_2 \mid C \subseteq U,\ C \notin \mathscr{D}_1\},\\
&\ \ \vdots\\
\mathscr{D}_k &= \{C \in \mathscr{C}_k \mid C \subseteq U,\ C \notin \mathscr{D}_1 \cup \cdots \cup \mathscr{D}_{k-1}\}\\
&\ \ \vdots
\end{aligned}$$

The result will follow if we can show that each point $\boldsymbol{x} \in U$ is contained in some $\mathscr{D}_k$, $k \in \mathbb{Z}_{>0}$. However, this follows since U is open, and so, for each $\boldsymbol{x} \in U$, one can find a smallest $k \in \mathbb{Z}_{\geq 0}$ with the property that there exists $C \in \mathscr{C}_k$ with $\boldsymbol{x} \in C$ and $C \subseteq U$. ▼

2 Lemma *The Lebesgue measure is the unique measure on* $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ *for which the measure of a rectangle is its volume.*

Proof From Theorem 5.4.2 we know that $\lambda_n(R) = \nu(R)$ for every rectangle R. Now suppose that $\mu\colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ is a measure with the property that $\mu(R) = \nu(R)$ for every rectangle R.

First let $U \subseteq \mathbb{R}^n$ be open. By Lemma 1 we can write $U = \cup_{j\in\mathbb{Z}_{>0}} C_j$ for a countable family $(C_j)_{j\in\mathbb{Z}_{>0}}$ of disjoint bounded cubes. Therefore, since μ is a measure,

$$\mu(U) = \mu\Big(\bigcup_{j\in\mathbb{Z}_{>0}} C_j\Big) = \sum_{j=1}^{\infty} \mu(C_j) = \sum_{j=1}^{\infty} \lambda_n(C_j) = \lambda_n(U).$$

Now let B be a bounded Borel set and let U be an open set for which $B \subseteq U$. Then

$$\mu(B) \leq \mu(U) = \lambda_n(U).$$

Therefore,

$$\mu(B) \leq \inf\{\lambda)n(U) \mid U \text{ open and } B \subseteq U\} = \lambda_n(B)$$

by regularity of λ_n. Therefore, if U is a bounded open set containing B we have

$$\mu(U) = \mu(B) + \mu(U \setminus B) \leq \lambda_n(B) + \lambda_n(U \setminus B) = \lambda_n(U).$$

Since $\mu(U) = \lambda_n(U)$ it follows that $\mu(B) = \lambda_n(B)$ and $\mu(U \setminus B) = \lambda_n(U \setminus B)$.

Finally let B be an unbounded Borel set. We can then write $B = \cup_{j_1,\dots,j_n\in\mathbb{Z}} B_{j_1\cdots j_n}$ where $(B_{j_1\cdots j_n})_{j_1,\dots,j_n\in\mathbb{Z}}$ is the (countable by Proposition **??**) family of disjoint Borel sets

$$B_{j_1\cdots j_n} = B \cap ([j_1, j_1+1) \times \cdots \times [j_1, j_1+1)), \qquad j_1, \dots, j_n \in \mathbb{Z}.$$

Then

$$\mu(B) = \sum_{j_1,\dots,j_n\in\mathbb{Z}} \mu(B_j) = \sum_{j_1,\dots,j_n\in\mathbb{Z}} \lambda_n(B_j) = \lambda_n(B),$$

as desired. ▼

To proceed with the proof, let $\mu\colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ be a translation-invariant measure and let $c = \mu([0,1)^n)$. By

$$\mu(\mathbb{R}^n) = \sum_{j_1,\dots,j_n\in\mathbb{Z}} \nu([j_1, j_1+1) \times \cdots \times [j_n, j_n+1)) = 0$$

by translation-invariance of μ. Now let $\mu'\colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ be the measure defined by $\mu'(B) = c^{-1}\mu(B)$. Now, for $k \in \mathbb{Z}_{\geq 0}$ let $\mathscr{C}_k$ be as in the proof of Lemma 1. Let $C \in \mathscr{C}_k$. We can write $[0,1)^n$ as a disjoint union of 2^{nk} intervals of the form $\boldsymbol{x}_j + C$. Therefore, by translation-invariance of μ',

$$\mu'([0,1)^n) = 2^{nk}\mu'(C), \quad \lambda_n([0,1)^n) = 2^{nk}\lambda_n(C).$$

Since $\mu'([0,1)^n) = \lambda_n([0,1)^n)$ it follows that $\mu'(C) = \lambda_n(C)$. Since every interval is a disjoint union of intervals from the sets $\mathscr{C}_k$, $k \in \mathbb{Z}_{\geq 0}$, by Lemma 1 it follows that $\mu'(C) = \lambda_n(C)$ for every cube C. Thus $\mu' = \lambda_n$ by Lemma 2 above and so $\mu = c\lambda_n$, as desired. ■

5.5.23 Theorem (There are no translation-invariant, length-preserving measures on all subsets of $\mathbb{R}^n$) *There exists no measure space $(\mathbb{R}^n, \mathscr{A}, \mu)$ having the joint properties that*

(i) $\mathscr{A} = 2^{\mathbb{R}^n}$,

(ii) $\mu((0,1)^n) = 1$, *and*

(iii) μ *is translation-invariant.*

Proof Were such a measure to exist, then the non-Lebesgue measurable set $A_n \subseteq (0,1)^n$ of Example 5.5.3 would be measurable. But during the course of Example 5.5.3 we saw that $(0,1)^n$ is a countable disjoint union of translates of A_n. The dichotomy illustrated in Example 5.5.3 then applies. That is, if $\mu(A_n) = 0$ then we get $\mu((0,1)^n) = 0$ and if $\mu(A_n) \in \mathbb{R}_{>0}$ then $\mu((0,1)^n) = \infty$, both of which conclusions are false. ■

Finally in this section, let us record another useful property of the Lebesgue measure, related to its being translation-invariant. From Definition **??** the notion of an orthogonal matrix, and the notation $\mathsf{O}(n)$ to denote the set of $n \times n$ orthogonal matrices.***missing stuff***

5.5.24 Definition (Rotation-invariant σ-algebra and measure on $\mathbb{R}^n$) A σ-algebra $\mathscr{A} \subseteq \mathbf{2}^{\mathbb{R}^n}$ is ***rotation-invariant*** if, for every $A \in \mathscr{A}$ and every $\boldsymbol{R} \in \mathsf{O}(n)$, $\boldsymbol{R}(A) \in \mathscr{A}$. A ***rotation-invariant*** measure on a rotation-invariant σ-algebra $\mathscr{A}$ is a map $\mu\colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ for which $\mu(\boldsymbol{R}(A)) = \mu(A)$ for every $A \in \mathscr{A}$ and $\boldsymbol{R} \in \mathsf{O}(n)$. •

We can then repeat the translation-invariant programme above for rotation-invariance. This begins with the following result.

5.5.25 Proposition (Rotation-invariance of the Borel and Lebesgue measurable sets) *Both $\mathscr{B}(\mathbb{R}^n)$ and $\mathscr{L}(\mathbb{R}^n)$ are rotation-invariant σ-algebras, and, moreover, λ_n is rotation invariant.*

Proof Let us denote

$$\mathscr{B}'(\mathbb{R}^n) = \{B \mid \boldsymbol{R}(B) \in \mathscr{B}(\mathbb{R}^n) \text{ for every } \boldsymbol{R} \in \mathsf{O}(n)\}.$$

We claim that $\mathscr{B}'(\mathbb{R}^n)$ is a σ-algebra containing the open subsets of $\mathbb{R}^n$. First of all, if $U \subseteq \mathbb{R}^n$ is open then $\boldsymbol{R}(U)$ is open for every $\boldsymbol{R} \in \mathsf{O}(n)$ since $\boldsymbol{R}$ is a homeomorphism of $\mathbb{R}^n$. Thus $U \in \mathscr{B}'(\mathbb{R}^n)$. To see that $\mathscr{B}'(\mathbb{R}^n)$ is a σ-algebra, first note that $\mathbb{R}^n = \boldsymbol{R}(\mathbb{R}^n)$ for every $\boldsymbol{R} \in \mathsf{O}(n)$ and so $\mathbb{R}^n \in \mathscr{B}'(\mathbb{R}^n)$. Next, let $B \in \mathscr{B}'(\mathbb{R}^n)$ and let $\boldsymbol{R} \in \mathsf{O}(n)$. Then

$$\begin{aligned}\boldsymbol{R}(\mathbb{R}^n \setminus B) &= \{\boldsymbol{R}(\boldsymbol{z}) \mid \boldsymbol{z} \notin B\} = \{\boldsymbol{y} \mid \boldsymbol{R}^{-1}(\boldsymbol{y}) \notin B\} = \{\boldsymbol{y} \mid \boldsymbol{y} \neq \boldsymbol{R}(\boldsymbol{z}),\ \boldsymbol{z} \in B\} \\ &= \{\boldsymbol{y} \mid \boldsymbol{y} \notin \boldsymbol{R}(B)\} = \mathbb{R}^n \setminus (\boldsymbol{R}(B)) \in \mathscr{B}(\mathbb{R}^n).\end{aligned}$$

Thus $\boldsymbol{R}(\mathbb{R}^n \setminus B) \in \mathscr{B}(\mathbb{R}^n)$ for every $\boldsymbol{R} \in \mathsf{O}(n)$ and so $\mathbb{R}^n \setminus B \in \mathscr{B}'(\mathbb{R}^n)$. Finally, let $(B_j)_{j\mathbb{Z}_{>0}}$ be a countable collection of subsets from $\mathscr{B}'(\mathbb{R}^n)$. Then, for $\boldsymbol{R} \in \mathsf{O}(n)$ we have

$$\boldsymbol{R}(\cup_{j\in\mathbb{Z}_{>0}} B_j) = \cup_{j\in\mathbb{Z}_{>0}} \boldsymbol{R}(B_j) \in \mathscr{B}(\mathbb{R}^n)$$

and so $\cup_{j\in\mathbb{Z}_{>0}} B_j \in \mathscr{B}'(\mathbb{R}^n)$. Thus $\mathscr{B}'(\mathbb{R}^n)$ is indeed a σ-algebra containing the open sets and so we conclude that $\mathscr{B}(\mathbb{R}^n) \subseteq \mathscr{B}'(\mathbb{R}^n)$ since $\mathscr{B}(\mathbb{R}^n)$ is the σ-algebra generated by the open sets. This shows that $\mathscr{B}(\mathbb{R}^n)$ is rotation-invariant.

Next let us show that $\mathscr{L}(\mathbb{R}^n)$ is rotation-invariant. To do this we first show that if $S \subseteq \mathbb{R}^n$ and if $\boldsymbol{R} \in \mathsf{O}(n)$ then $\lambda_n^*(\boldsymbol{R}(S)) = \lambda_n^*(S)$. First note by Theorem **??** that $\nu(\boldsymbol{R}(R)) = \nu(R)$ since $\det \boldsymbol{R} \in \{-1, 1\}$ (see Exercise **??**). Then we compute

$$\begin{aligned}\lambda_n^*(\boldsymbol{R}(S)) &= \inf\Big\{\sum_{j=1}^{\infty} \nu(R_j)\ \Big|\ \boldsymbol{R}(S) \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} R_j\Big\} \\ &= \inf\Big\{\sum_{j=1}^{\infty} \nu(R'_j)\ \Big|\ \boldsymbol{R}(S) \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} \nu(\boldsymbol{R}(R'_j))\Big\} \\ &= \inf\Big\{\sum_{j=1}^{\infty} \nu(R'_j)\ \Big|\ S \subseteq \bigcup_{j\in\mathbb{Z}_{>0}} R'_j\Big\} = \lambda_n^*(S).\end{aligned}$$

Now let $A \in \mathscr{L}(\mathbb{R}^n)$ so that, for every subset $S \subseteq \mathbb{R}^n$,

$$\lambda_n^*(S) = \lambda_n^*(S \cap A) + \lambda_n^*(S \cap (\mathbb{R}^n \setminus A)).$$

Then, for $\boldsymbol{R} \in \mathsf{O}(n)$ and $S \subseteq \mathbb{R}^n$,

$$\lambda_n^*(S \cap (\boldsymbol{R}(A))) = \lambda_n^*((\boldsymbol{R}\boldsymbol{R}^{-1}(S)) \cap (\boldsymbol{R}(A))) = \lambda_n^*((\boldsymbol{R}^{-1}(S)) \cap A)$$

and, similarly,

$$\lambda_n^*(S \cap (\mathbb{R}^n \setminus (\boldsymbol{R}(A)))) = \lambda_n^*((\boldsymbol{R}\boldsymbol{R}^{-1}(S)) \cap (\boldsymbol{R}(\mathbb{R}^n \setminus A))) = \lambda_n^*((\boldsymbol{R}^{-1}(S)) \cap (\mathbb{R}^n \setminus A)).$$

Since $\lambda_n^*(\boldsymbol{R}^{-1}(S)) = \lambda_n^*(S)$ this immediately gives

$$\lambda_n^*(S) = \lambda_n^*(S \cap (\boldsymbol{R}(A))) + \lambda_n^*(S \cap (\mathbb{R}^n \setminus (\boldsymbol{R}(A)))),$$

showing that $\boldsymbol{R}(A) \in \mathscr{L}(\mathbb{R}^n)$.

The final assertion in the statement of the result, that λ_n is rotation-invariant, follows from the fact, proved above, that $\lambda_n^*(S) = \lambda_n^*(\boldsymbol{R}(S))$ for every $S \subseteq \mathbb{R}^n$. ■

The following generalisation of the preceding result is also useful.

5.5.26 Proposition (Lebesgue measure and linear maps) *If* $\mathbf{L} \in \mathrm{L}(\mathbb{R}^n;\mathbb{R}^m)$ *then* $\mathrm{matL(B)} \in \mathscr{B}(\mathbb{R}^m)$ *if* $\mathrm{B} \in \mathscr{B}(\mathbb{R}^n)$ *and* $\mathbf{L}(\mathrm{A}) \in \mathscr{L}(\mathbb{R}^m)$ *if* $\mathrm{A} \in \mathscr{L}(\mathbb{R}^n)$. *Moreover, if* $\mathrm{A} \in \mathscr{L}(\mathbb{R}^n)$ *then* $\lambda_{\mathrm{m}}(\mathbf{L}(\mathrm{A})) = \det \mathbf{L}\lambda_{\mathrm{n}}(\mathrm{A})$.

Proof ■

5.5.4 Lebesgue measure on $\mathbb{R}^n$ as a product measure

The Lebesgue measure on $\mathbb{R}^n$ is *not* the product of the Lebesgue measures on the factors of $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$. The problem, as we shall see, is that the product of the Lebesgue measures is not complete. Fortunately, while the Principle of Desired Commutativity does not apply in its simplest form, it is not too far off since the Lebesgue measure on $\mathbb{R}^n$ is the *completion* of the product measure.

First we consider the relationship between the measure spaces $(\mathbb{R}^n, \sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})), \lambda \times \cdots \times \lambda)$ and $(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n), \lambda_n)$. The first observation is the following.

5.5.27 Proposition ($\sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})) \subseteq \mathscr{L}(\mathbb{R}^n)$) *We have* $\sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})) \subseteq \mathscr{L}(\mathbb{R}^n)$.

Proof By definition, $\sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R}))$ is the σ-algebra generated by the measurable rectangles in $\mathbb{R}^n$. It, therefore, suffices to show that measurable rectangles are λ_n-measurable. Thus let $A_1 \times \cdots \times A_n$ be a measurable rectangle and, by Corollary 5.4.12, write $A_j = B_j \cup Z_j$ for $B_j \in \mathscr{B}(\mathbb{R})$ and $Z_j \subseteq \mathbb{R}$ having measure zero. Then $A_1 \times \cdots \times A_n$ is a union of measurable rectangles of the form $S_1 \times \cdots \times S_n$ where $S_j \in \{B_j, Z_j\}$, $j \in \{1, \ldots, n\}$. We claim that if $S_j = Z_j$ for some $j \in \{1, \ldots, n\}$ then the corresponding measurable rectangle has Lebesgue measure zero, and so in particular is Lebesgue measurable. To see this, consider a measurable rectangle of the form

$$S_1 \times \cdots \times S_{j-1} \times Z_j \times S_{j+1} \times \cdots \times S_n.$$

Let $k \in \mathbb{Z}_{>0}$ and let $C_k = [-k, k]$. Let $\epsilon \in \mathbb{R}_{>0}$. Since Z_j has measure zero, there exists intervals (a_l, b_l), $l \in \mathbb{Z}_{>0}$, such that $Z_j \subseteq \cup_{l \in \mathbb{Z}_{>0}}(a_l, b_l)$ and

$$\sum_{l=1}^{\infty}(b_l - a_l) < \frac{\epsilon}{(2k)^{n-1}}.$$

Therefore,

$$\lambda_n(C_k \cap (S_1 \times \cdots \times S_{j-1} \times Z_j \times S_{j+1} \times \cdots \times S_n)) < (2k)^{n-1}\frac{\epsilon}{(2k)^{n-1}} = \epsilon.$$

Thus

$$\lambda_n(C_k \cap (S_1 \times \cdots \times S_{j-1} \times Z_j \times S_{j+1} \times \cdots \times S_n)) = 0$$

and, since

$$S_1 \times \cdots \times S_{j-1} \times Z_j \times S_{j+1} \times \cdots \times S_n \\ = \cup_{k\in\mathbb{Z}_{>0}}(C_k \cap (S_1 \times \cdots \times S_{j-1} \times Z_j \times S_{j+1} \times \cdots \times S_n)),$$

it follows from Proposition 5.3.3 that

$$\lambda_n(S_1 \times \cdots \times S_{j-1} \times Z_j \times S_{j+1} \times \cdots \times S_n) = 0,$$

as desired. Thus the only measurable rectangle comprising $A_1\times\cdots\times A_n$ that is possibly not of measure zero is $B_1\times\cdots\times B_n$. By Proposition 5.5.7 (and its natural generalisation to more than two factors using a trivial induction) it follows that this set will be Borel measurable, and so Lebesgue measurable. Thus $A_1 \times \cdots \times A_n$ is a finite union of Lebesgue measurable sets and so is Lebesgue measurable. ■

An example illustrates that the inclusion from the preceding proposition is strict.

5.5.28 Example ($\sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})) \subset \mathscr{L}(\mathbb{R}^n)$) Let

$$A = \mathbb{R} \times \{\mathbf{0}_{n-1}\} \subseteq \mathbb{R}^n$$

and note that by Theorem 5.3.33 we have $\lambda \times \cdots \times \lambda(A) = 0$. Now let $E \subseteq \mathbb{R}$ be a non-Lebesgue measurable set and note that $S \triangleq E \times \{\mathbf{0}_{n-1}\} \subseteq A$ is thus a subset of measure zero, and thus an element of $\mathscr{L}(\mathbb{R}^n)$ by completeness of the n-dimensional Lebesgue measure. We claim that $S \notin \sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R}))$. Indeed, by Proposition 5.2.18 it follows that if S is measurable then E must be measurable, which it is not. •

Thus the Lebesgue measure on $\mathbb{R}^n$ is *not* the product of the Lebesgue measures on its $\mathbb{R}$ factors. However, all is not lost, as the following result suggests.

5.5.29 Proposition (The Lebesgue measure is the completion of the product measure) *The measure space* $(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n), \lambda_n)$ *is the completion of the measure space* $(\mathbb{R}^n, \sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})), \lambda \times \cdots \times \lambda)$.

Proof Note that open rectangles are in $\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})$. Thus, since $\mathscr{B}(\mathbb{R}^n)$ is the σ-algebra generated by open rectangles, $\mathscr{B}(\mathbb{R}^n) \subseteq \sigma(\mathscr{L}(\mathbb{R})\times\cdots\times\mathscr{L}(\mathbb{R}))$. Moreover, for an open rectangle we have $U_1 \times \cdots \times U_n$ we have

$$\lambda \times \cdots \times \lambda(U_1 \times \cdots \times U_n) = \lambda_n(U_1 \times \cdots \times U_n).$$

By Lemma 2 of Theorem 5.5.22 we then have

$$\lambda \times \cdots \times \lambda|\mathscr{B}(\mathbb{R}^n) = \lambda_n|\mathscr{B}(\mathbb{R}^n).$$

Now, by Proposition 5.5.27, we have

$$\mathscr{B}(\mathbb{R}^n) \subseteq \sigma(\mathscr{L}(\mathbb{R}) \times \cdots \times \mathscr{L}(\mathbb{R})) \subseteq \mathscr{L}(\mathbb{R}^n)$$

with $\lambda \times \cdots \times \lambda$ and λ_n agreeing on the left and right sets. By Theorem 5.5.12 the result follows. ■

5.5.5 Coverings of subsets of $\mathbb{R}^n$

It is useful to sometimes be able to cover subsets of $\mathbb{R}^n$ with certain types of sets—say, open balls—and such that the covering has certain desired properties. In this section we give a few such results that are useful and some of which are related to the Lebesgue measure. Various versions of the results here are known as the ***Vitali***[6] ***Covering Lemma***.

The most basic such result, and the starting point for other results, is the following.

5.5.30 Lemma (Properties of coverings by balls) *Let J be an index set and let* $(\overline{\mathsf{B}}^n(r_j, \mathbf{x}_j))_{j\in\mathbb{Z}_{>0}}$ *be a family of balls such that*

$$\sup\{r_j \mid j \in J\} < \infty.$$

Then there exists a subset $J' \subseteq J$ *with the following properties:*

(i) J' *is finite or countable;*

(ii) the balls $(\overline{\mathsf{B}}^n(r_{j'}, \mathbf{x}_{j'}))_{j'\in J'}$ *are pairwise disjoint;*

(iii) $\cup_{j\in J}\overline{\mathsf{B}}^n(r_j, \mathbf{x}_j) \subseteq \cup_{j'\in J'}\overline{\mathsf{B}}^n(5r_{j'}, \mathbf{x}_{j'})$.

Proof Let us first suppose that $\cup_{j\in J}\overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ is bounded. We inductively construct a subset J' of J as follows. Let $\rho_1 = \sup\{r_j \mid j \in J\}$ and let $j_1 \in J$ be chosen so that $r_j \geq \frac{1}{2}\rho_1$. Now suppose that $j_1, \dots, j_k$ have been defined and let

$$\rho_{k+1} = \sup\{r_j \mid j \text{ satisfies } \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap \cup_{s=1}^k \overline{\mathsf{B}}^n(r_{j_s}, \boldsymbol{x}_{j_s}) = \emptyset\}.$$

If $\rho_{k+1} = 0$ then take $J' = \{j_1, \dots, j_k\}$. Otherwise define $j_{k+1} \in J \setminus \{j_1, \dots, j_k\}$ such that $r_{j_{k+1}} \geq \frac{1}{2}\rho_{k+1}$. In the case where this inductive procedure does not terminate in finitely many steps, take $J' = \{j_k \mid k \in \mathbb{Z}_{>0}\}$.

The family $(\overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'}))_{j'\in J'}$ so constructed is clearly pairwise disjoint. Moreover, if $\boldsymbol{x} \in \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ for some $j \in J$ we have two possibilities.

1. $j \in J'$: In this case we immediately have $\boldsymbol{x} \in \cup_{j'\in J'}\overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'})$.
2. $j \notin J'$: Here we claim that there exists $j' \in J$ such that $\boldsymbol{x} \in \overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'})$. Suppose otherwise. Note that since we are assuming that $\cup_{j\in J}\overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ is bounded and since $(\overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'}))_{j'\in J'}$ is pairwise disjoint, for every $\epsilon \in \mathbb{R}_{>0}$ we have $\lim_{k\to\infty} r_{j_k} = 0$. Therefore, there must exist $k \in \mathbb{Z}_{>0}$ such that $2r_{j_k} < r_j$. This, however, contradicts the definition of r_k, and so we must have $\boldsymbol{x} \in \overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'})$ for some $j' \in J'$.

To complete the proof in this case we prove a simple geometrical lemma.

[6] Giuseppe Vitali (1875–1932) was an Italian Mathematician who made important contributions to analysis.

1 Sublemma *Let* $x_1, x_2 \in \mathbb{R}^n$*, let* $r_1, r_2 \in \mathbb{R}_{>0}$ *satisfy* $r_1 \geq \frac{1}{2} r_2$*, and suppose that* $\overline{\mathsf{B}}^n(r_1, x_1) \cap \overline{\mathsf{B}}^n(r_2, x_2) \neq \emptyset$*. Then* $\overline{\mathsf{B}}^n(r_2, x_2) \subseteq \overline{\mathsf{B}}^n(5r_1, x_1)$*.*

Proof Let $\boldsymbol{x} \in \overline{\mathsf{B}}^n(r_2, \boldsymbol{x}_2)$ and let $\boldsymbol{y} \in \overline{\mathsf{B}}^n(r_1, \boldsymbol{x}_1) \cap \overline{\mathsf{B}}^n(r_2, \boldsymbol{x}_2)$. Multiple applications of the triangle inequality gives

$$\|\boldsymbol{x} - \boldsymbol{x}_1\|_{\mathbb{R}^n} \leq \|\boldsymbol{x} - \boldsymbol{x}_2\|_{\mathbb{R}^n} + \|\boldsymbol{y} - \boldsymbol{x}_2\|_{\mathbb{R}^n} + \|\boldsymbol{y} - \boldsymbol{x}_1\|_{\mathbb{R}^n} \leq 5r_1,$$

as desired. ▼

From the sublemma and since we have shown that, for each $j \in J$, $\overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ intersects at least one of the balls $\overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'})$, $j' \in J'$, it follows that

$$\cup_{j \in J} \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \subseteq \cup_{j' \in J'} \overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'}),$$

as claimed.

Next we consider the case where $\cup_{j \in J} \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ is not bounded. Let $\rho = \sup\{r_j \mid j \in J\}$. We inductively define J'_k, $k \in \mathbb{Z}_{>0}$, of J as follows. Define

$$J_1 = \{j \in J \mid \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap \overline{\mathsf{B}}^n(4\rho, \boldsymbol{0}_n) \neq \emptyset\}$$

and note that $\cup_{j \in J_1} \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ is bounded since ρ is finite. Let $J''_1 \subseteq J_1$ be defined by the applying the procedure from the first part of the proof to the set J_1. Then denote

$$J'_1 = \{j \in J''_1 \mid \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap \overline{\mathsf{B}}^n(\rho, \boldsymbol{0}_n) \neq \emptyset\}.$$

Note that

1. $(\overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'}))_{j' \in J'_1}$ are pairwise disjoint and that
2. $\cup\{\overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \mid \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap \overline{\mathsf{B}}^n(\rho, \boldsymbol{0}_n)\} \subseteq \cup_{j' \in J'_1} \overline{\mathsf{B}}^n(5r_{j'}, \boldsymbol{x}_{j'})$.

Next define

$$J_2 = J_1 \cup \{j \in J \mid \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap (\overline{\mathsf{B}}^n(5\rho, \boldsymbol{0}_n) \setminus \overline{\mathsf{B}}^n(4\rho, \boldsymbol{0}_n)) \neq \emptyset\}.$$

Also take $J''_2 \subseteq J_2$ to be the subset constructed as in the first part of the proof. Then define

$$J'_2 = \{j \in J''_2 \mid \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap \overline{\mathsf{B}}^n(2\rho, \boldsymbol{0}_n) \neq \emptyset\}.$$

Note that since the only balls added to J_1 in forming J_2 do not intersect $\overline{\mathsf{B}}^n(\rho, \boldsymbol{0}_n)$, it follows that $J''_1 \subseteq J''_2$, and thus that $J'_1 \subseteq J'_2$. Moreover, note that

1. $(\overline{\mathsf{B}}^n(r_{j'}, \boldsymbol{x}_{j'}))_{j' \in J'_2}$ are pairwise disjoint and that
2. $\cup\{\overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \mid \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j) \cap \overline{\mathsf{B}}^n(2\rho, \boldsymbol{0}_n)\} \subseteq \cup_{j' \in J'_2} \overline{\mathsf{B}}^n(5r_{j'}, \boldsymbol{x}_{j'})$.

Proceeding in this way we define $J'_1 \subseteq \cdots \subseteq J'_k \subseteq \cdots$. Then take $J' = \cup_{k \in \mathbb{Z}_{>0}} J'_k$. By Proposition **??** it follows that J' is countable. If $j'_1, j'_2 \in J'$ then, by construction, there exists $k \in \mathbb{Z}_{>0}$ such that $j'_1, j'_2 \in J'_k$. It thus follows that $\overline{\mathsf{B}}^n(r_{j'_1}, \boldsymbol{x}_{j'_1})$ and $\overline{\mathsf{B}}^n(r_{j'_2}, \boldsymbol{x}_{j'_2})$ are disjoint. If $\boldsymbol{x} \in \cup_{j \in J} \overline{\mathsf{B}}^n(r_j, \boldsymbol{x}_j)$ we have $\boldsymbol{x} \in \overline{\mathsf{B}}^n(r_{j_0}, \boldsymbol{x}_{j_0})$ where $\overline{\mathsf{B}}^n(r_{j_0}, \boldsymbol{x}_{j_0}) \cap \overline{\mathsf{B}}^n(k\rho, \boldsymbol{0}_m) \neq \emptyset$ for some $k \in \mathbb{Z}_{>0}$. Then $\boldsymbol{x} \in \cup_{j' \in J'_k} \overline{\mathsf{B}}^n(5r_{j'}, \boldsymbol{x}_{j'})$. This gives the lemma. ■

5.5.31 Remark (The Vitali Covering Lemma for finite coverings) If the index set J is finite in the preceding result, then one can strengthen the conclusions to assert that

$$\cup_{j\in J}\overline{\mathsf{B}}^n(r_j,\boldsymbol{x}_j)\subseteq\cup_{j'\in J'}\overline{\mathsf{B}}^n(3r_{j'},\boldsymbol{x}_{j'}).$$

This is achieved merely by noting that one can choose the numbers $j_1,\dots,j_k$ so that $r_s=\rho_s$ for $s\in\{1,\dots,k\}$. In this case, the factor of $\frac{1}{2}$ in the sublemma can be removed with the resulting change of the factor 5 to 3. •

There is a similar such result for, not balls, but cubes. Recall from Section **??** that a cube in $\mathbb{R}^n$ is a rectangle, all of whose sides have the same length. We shall denote by

$$\overline{\mathsf{C}}(r,\boldsymbol{x})=[x_1-r,x_1+r]\times\dots\times[x_n-r,x_n+r]$$

the cube centred at $\boldsymbol{x}\in\mathbb{R}^n$ and with sides of length $2r$.

5.5.32 Lemma (Properties of coverings by cubes) *Let* $\mathrm{J}\in\{\{1,\dots,\mathrm{m}\},\mathbb{Z}_{>0}\}$ *and let* $(\overline{\mathsf{C}}(\mathrm{r_j},\mathbf{x_j}))_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be a finite or countable family of cubes. Then there exists a subset* $\mathrm{J}'\subseteq\mathrm{J}$ *with the following properties:*

(i) the cubes $(\overline{\mathsf{C}}(\mathrm{r_{j'}},\mathbf{x_{j'}}))_{\mathrm{j'}\in\mathrm{J'}}$ *are pairwise disjoint;*

(ii) $\cup_{\mathrm{j}\in\mathrm{J}}\overline{\mathsf{C}}(\mathrm{r_j},\mathbf{x_j})\subseteq\cup_{\mathrm{j'}\in\mathrm{J'}}\overline{\mathsf{C}}(5\mathrm{r_{j'}},\mathbf{x_{j'}})$.

Proof The result follows easily after making some observations about cubes, relying on the general notion of a norm that we will introduce and discuss in Section 6.1. If we define

$$\|(x_1,\dots,x_n)\|_\infty=\max\{|x_1|,\dots,|x_n|\},$$

then we note from Example 6.1.3–4 that this defines a norm on $\mathbb{R}^n$. Moreover, the balls in this norm are cubes, as can be easily verified. A review of the proof of Lemma 5.5.30 shows that it is the norm properties of $\|\cdot\|_{\mathbb{R}^n}$ that are used, along with the fact that the balls in the norm $\|\cdot\|_{\mathbb{R}^n}$ are the balls in the usual sense. Thus the entire proof of Lemma 5.5.30 carries over, replacing $\overline{\mathsf{B}}^n(r,\boldsymbol{x})$ with $\overline{\mathsf{C}}(r,\boldsymbol{x})$ and $\|\cdot\|_{\mathbb{R}^n}$ with $\|\cdot\|_\infty$. ■

The importance of the preceding results is not so readily seen at a first glance. To illustrate the essence of the result, consider the following observation. Using the notation of Lemma 5.5.30, suppose that $\cup_{j\in J}\overline{\mathsf{B}}^n(r_j,\boldsymbol{x}_j)$ is bounded. The preceding result says that there is a countable *disjoint* subset these balls that covers at least $\frac{1}{5^n}$ of the volume of region covered by the complete collection of balls. The main point is that the volume fraction covered by the disjoint balls is bounded below by a quantity, $\frac{1}{5^n}$, that is independent of the covering. It is this property that will make these preceding lemmata useful in the subsequent discussion.

First we give a definition for a sort of covering which we shall show has useful properties. In the following definition, let us call the number $2r$ the ***diameter*** of a ball $\overline{\mathsf{B}}^n(r,\boldsymbol{x})$ or a cube $\overline{\mathsf{C}}(r,\boldsymbol{x})$. We denote the diameter of a ball or cube B by $\operatorname{diam}(B)$.

5.5.33 Definition (Vitali covering) Let J be an index set, let $A \subseteq \mathbb{R}^n$, and let $(B_j)_{j\in J}$ be a family of either closed balls or closed cubes, i.e., either (1) B_j is a closed ball for every $j \in J$ or (2) B_j is a closed cube for every $j \in J$. Suppose that $\operatorname{int}(B_j) \neq \emptyset$ for each $j \in J$. This family of balls or cubes is a ***Vitali covering*** of A if, for every $\epsilon \in \mathbb{R}_{>0}$ and for every $\boldsymbol{x} \in A$ there exists $j \in J$ such that $\boldsymbol{x} \in B_j$ and such that $\operatorname{diam}(B_j) < \epsilon$. •

Before giving the essential theorem about Vitali coverings, let us give an example of a Vitali covering.

5.5.34 Example (Vitali covering) If $A \subseteq \mathbb{R}^n$, define

$$\mathscr{C}_A = \{\overline{\mathsf{B}}^n(r, \boldsymbol{x}) \subseteq \mathbb{R}^n \mid r \in \mathbb{R}_{>0},\ A \cap \overline{\mathsf{B}}^n(r, \boldsymbol{x}) \neq \emptyset\}.$$

If $\epsilon \in \mathbb{R}_{>0}$ and $\boldsymbol{x} \in A$ then $\overline{\mathsf{B}}^n(\epsilon, \boldsymbol{x})$ contains $\boldsymbol{x}$ and is in $\mathscr{C}_A$. This implies that $\mathscr{C}_A$ is a Vitali covering. •

As the definition implies and the above example illustrates, one might expect that a Vitali covering of a set will involve a plentiful, rather than a barely sufficient, collection of balls or cubes.

The following theorem will be useful for us in a few different places in the text.

5.5.35 Theorem (Property of Vitali coverings) *Let* $\mathrm{A} \subseteq \mathrm{R}^n$ *be nonempty and let* $(\mathrm{B_j})_{\mathrm{j}\in\mathrm{J}}$ *be a Vitali covering of* A *by cubes or balls. Then there exists a countable or finite subset* $\mathrm{J}' \subseteq \mathrm{J}$ *such that*

(i) the sets $(\mathrm{B_{j'}})_{\mathrm{j'}\in\mathrm{J'}}$ *are pairwise disjoint and*

(ii) $\lambda_{\mathrm{n}}^*(\mathrm{A} - \cup_{\mathrm{j'}\in\mathrm{J}}\mathrm{B_{j'}}) = 0.$

Proof First we suppose that A is bounded and let U be a bounded open set such that $A \subseteq U$ and define

$$J'' = \{j \in J \mid B_j \subseteq U\}$$

and note that $(B_j)_{j\in J''}$ is a Vitali cover of A (why?). We now apply the construction of either of Lemma 5.5.30 or 5.5.32 as appropriate to arrive at a finite or countable subset $J' \subseteq J''$. For the remainder of the proof, for concreteness let us suppose that J' is infinite and write $J' = \{j_k\}_{k\in\mathbb{Z}_{>0}}$. We also recall from the proof of Lemma 5.5.30 the sequence $(\rho_k)_{k\in\mathbb{Z}_{>0}}$ of positive numbers.

Now let $N \in \mathbb{Z}_{>0}$ and let $\boldsymbol{x} \in A - \cup_{k=1}^N B_{j_k}$. Since the set $\cup_{k=1}^N B_{j_k}$ is closed by Proposition **??** and since $(B_j)_{j\in J''}$ is a Vitali covering of A, there exists $j \in J''$ such that $\boldsymbol{x} \in B_j$ and $B_j \cap (\cup_{k=1}^N B_{j_k}) = \emptyset$. Suppose $m \in \mathbb{Z}_{>0}$ is such that $B_j \cap (\cup_{k=1}^m B_{j_k} = \emptyset$. Then $\operatorname{diam}(B_j) \leq \rho_{k+1}$ by definition of ρ_{k+1}. Since $\lim_{k\to\infty} \rho_k = 0$ (see the proof of Lemma 5.5.30) it must therefore be the case that there exists $m_0 \in \mathbb{Z}_{>0}$ such that $B_j \cap (\cup_{k=1}^m B_{j_k}) \neq \emptyset$ for all $m \geq m_0$. Thus $\operatorname{diam}(B_j) \leq \rho_{m_0}$ and so $\operatorname{diam}(B_j) \leq 2\operatorname{diam}(B_{m_0})$ since $\operatorname{diam}(B_{m_0}) \geq \frac{1}{2}\rho_{m_0}$. Since $B_j \cap (\cup_{k=1}^{m_0-1} B_{j_k}) = \emptyset$ we must have $B_j \cap C_{m_0} \neq \emptyset$. For $j \in J$ let B'_j be the ball or cube whose centre agrees with that of B_j but for which $\operatorname{diam}(B'_j) = 5\operatorname{diam}(B_j)$. The lemma from the proof of Lemma 5.5.30 then gives $B_j \subseteq B'_{m_0}$. Since $m_0 \geq N + 1$ by virtue of the fact that $B_j \cap (\cup_{k=1}^N B_{j_k}) = \emptyset$, we then have

$$\boldsymbol{x} \in B_j \subseteq B'_{m_0} \subseteq \cup_{k=N+1} B'_{j_k}.$$

This shows that

$$A - \cup_{k=1}^{N} B_{j_k} \subseteq \cup_{k=N+1} B'_{j_k}.$$

Now note that $\sum_{k=1}^{\infty} \lambda_n(B_{j_k}) < \infty$, as was shown during the proof of Lemma 5.5.30. An application of Exercise 5.5.1 then gives $\sum_{k=1}^{\infty} \lambda_n(B'_{j_k}) < \infty$. Let $\epsilon \in \mathbb{R}_{>0}$. By Proposition 2.4.7 it follows that there exists $N \in \mathbb{Z}_{>0}$ sufficiently large that $\sum_{k=N+1}^{\infty} \lambda_n(B'_{j_k}) < \epsilon$. Therefore,

$$\lambda_n^*\Big(A - \cup_{k=1}^{N} B_{j_k}\Big) \leq \lambda_n^*\Big(\cup_{k=N+1} B'_{j_k}\Big) = \sum_{k=N+1}^{\infty} \lambda_n^*(B'_{j_k}) < \epsilon,$$

using monotonicity and subadditivity of the Lebesgue outer measure. Monotonicity of the Lebesgue outer measure shows that

$$\lambda_n^*\Big(A - \cup_{k=1}^{\infty} B_{j_k}\Big) \leq \lambda_n^*\Big(A - \cup_{k=1}^{N} B_{j_k}\Big) < \epsilon,$$

which completes the proof in the case that A is bounded.

If A is unbounded, proceed as follows. Let $(U_k)_{k \in \mathbb{Z}_{>0}}$ be a countable collection of pairwise disjoint bounded open sets for which

$$\lambda_n(\mathbb{R}^n \setminus \cup_{k=1}^{\infty} U_k) = 0.$$

Let $A_k = U_k \cap A$. For every $k \in \mathbb{Z}_{>0}$ for which $A_k \neq \emptyset$ the first part of the proof yields a finite or countable subset $J'_k \subseteq J$ such that the family $(B_{j'_k})_{j'_k \in J'_K}$ is pairwise disjoint and such that

$$\lambda_n^*\Big(A_k - \cup_{j'_k \in J'_k} B_{j'_k}\Big) = 0.$$

Let us define $J' = \cup_{k=1}^{\infty} J'_k$ and note that, by virtue of the constructions in the first part of the proof, $(B_{j'})_{j' \in J'}$ is pairwise disjoint. Moreover,

$$A = \cup_{k=1}^{\infty} A_k \cup (A \cap (\mathbb{R}^n \setminus \cup_{l=1}^{\infty} U_l))$$

from which we conclude that

$$A - \cup_{j' \in J'} B_{j'} = (\cup_{k=1}^{\infty} A_k - \cup_{j_k=1}^{\infty} B_{j'_k}).$$

Note that J' is countable by Proposition **??**. Thus $A - \cup_{j' \in J'} B_{j'}$ is a countable union of sets of measure zero, and so is a set of measure zero. ■

5.5.6 The Banach–Tarski Paradox

In this section we give an "elementary" proof of an (in)famous result regarding the strangeness of sets that are not Lebesgue measurable. Let us state the result first and then provide some discussion. After this we will devote the remainder of the section to the proof of the theorem.

To state the result we first introduce some language to organise the statement. We recall from Definition **??** the definition of an isometry and from Theorem **??** the characterisation of characterisation of isometries. The group of isometries is denoted in Definition **??** by $\mathsf{E}(n)$.

5.5.36 Definition (Piecewise congruent) Subsets $X, Y \subseteq \mathbb{R}^n$ are ***piecewise congruent*** if there exists

(i) $N \in \mathbb{Z}_{>0}$,

(ii) a partition $(X_1, \dots, X_N)$ of X, and

(iii) $\rho_1, \dots, \rho_N \in \mathsf{E}(n)$

such that $(\rho_1(X_1), \dots, \rho_N(X_N))$ is a partition of Y. •

Piecewise congruence should be viewed as follows. The set X is chopped up into N bits, and these bits are rearranged without distortion to give Y. An illustration of this in a simple case is given in Figure 5.4. The idea seems innocuous enough, but the Banach–Tarski Paradox tells us that some unexpected sets can be piecewise congruent.

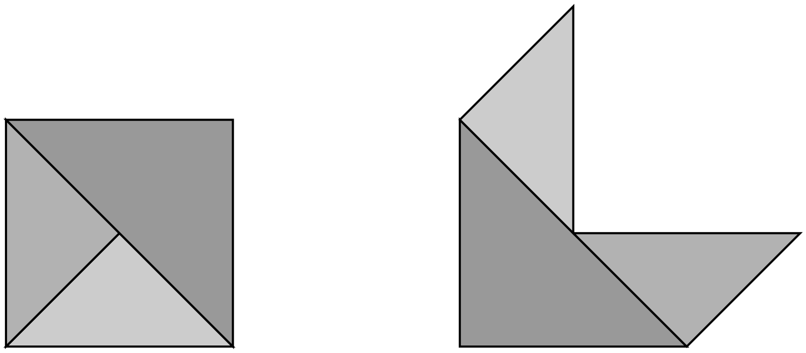


Figure 5.4 Piecewise congruent sets

5.5.37 Theorem (Banach–Tarski Paradox) *If* X, Y $\subseteq \mathbb{R}^3$ *are bounded sets with nonempty interiors then they are piecewise congruent.*

For example, the result says that one can cut up a set the size a pea into a finite number of disjoint components and reassemble these into a set the size of Jupiter. A common first reaction to this is that it is obviously false. But one should take care to understand that the theorem does not say this is true in the physical world, only in the mathematical world. In the mathematical world, or at least the one with the Axiom of Choice, there are sets whose volume does not behave as one normally expects volume to behave. It is this sort of set into which the set X is being partitioned in the theorem. For example, one should consider the set A of Example 5.4.3 that is not Lebesgue measurable. The main idea in showing that A is not Lebesgue measurable consists of showing that $(0, 1)$ can be written as a *countable* disjoint union of translates of A. This led us directly to contradictory conclusions that, if the volume of A is well-behaved, then $(0, 1)$ has either zero or infinite volume. Well, the subsets into which the sets of the Banach–Tarski Paradox are partitioned are non-Lebesgue measurable too. Thus we should not expect that the volumes of these sets behave in a decent way.

Let us now prove the Banach–Tarski Paradox. The proof is involved, but elementary. In the proof we denote by $\mathbb{S}^2$ the boundary of $\overline{\mathsf{B}}^3(1,\mathbf{0})$, i.e., $\mathbb{S}^2$ is the sphere of radius 1 in $\mathbb{R}^3$.

Our proof begins with some algebraic constructions. Define $\boldsymbol{A},\boldsymbol{B}\in\mathsf{O}(3)$ by

$$\boldsymbol{A}=\begin{bmatrix}-\cos\theta & 0 & \sin\theta\\ 0 & -1 & 0\\ \sin\theta & 0 & \cos\theta\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0\\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0\\ 0 & 0 & 1\end{bmatrix}.$$

The value of $\theta\in\mathbb{R}$ will be chosen shortly. One verifies directly that

$$\boldsymbol{B}^2=\begin{bmatrix}-\frac{1}{2} & \frac{\sqrt{3}}{2} & 0\\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0\\ 0 & 0 & 1\end{bmatrix},\quad \boldsymbol{B}^3=\boldsymbol{A}^2=\boldsymbol{I}_3.$$

Thus $\boldsymbol{A}^{-1}=\boldsymbol{A}$ and $\boldsymbol{B}^{-1}=\boldsymbol{B}^2$. It then follows that if we define

$$\mathsf{G}=\{\boldsymbol{R}_1\cdots\boldsymbol{R}_k \mid k\in\mathbb{Z}_{>0},\ \boldsymbol{R}_j\in\{\boldsymbol{A},\boldsymbol{B}\},\ j\in\{1,\dots,k\}\}\cup\{\boldsymbol{I}_3\},$$

then G is a subgroup of $\mathsf{O}(3)$. Note that it is possible that

$$\boldsymbol{R}_1\cdots\boldsymbol{R}_k=\boldsymbol{R}'_1,\dots,\boldsymbol{R}'_{k'},$$

i.e., that different products will actually agree. We wish to eliminate this ambiguity. First of all, note that the relations $\boldsymbol{A}^3=\boldsymbol{B}^2=\boldsymbol{I}_3$ ensure that if $\boldsymbol{R}_1,\dots,\boldsymbol{R}_k\in\{\boldsymbol{A},\boldsymbol{B}\}$ then we can write

$$\boldsymbol{R}_1\cdots\boldsymbol{R}_k=\boldsymbol{R}'_1\cdots\boldsymbol{R}'_{k'}$$

for $\boldsymbol{R}'_j\in\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{B}^2\}$, $j\in\{1,\dots,k'\}$. Next we claim that if $\boldsymbol{R}_1,\dots,\boldsymbol{R}_k\in\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{B}^2\}$ for $k\geq 2$ then

$$\boldsymbol{R}\triangleq\boldsymbol{R}_1\cdots\boldsymbol{R}_k=\boldsymbol{R}'_1\cdots\boldsymbol{R}'_k$$

where $\boldsymbol{R}'_j\in\{\boldsymbol{B}\boldsymbol{A},\boldsymbol{B}^2\boldsymbol{A}\}$. This, however, follows from the fact that the relations $\boldsymbol{A}^3=\boldsymbol{B}^2=\boldsymbol{I}_3$ ensure that at least one of the following four possibilities hold:

1. $\boldsymbol{R}=\boldsymbol{B}^{r_1}\boldsymbol{A}\boldsymbol{B}^{r_2}\boldsymbol{A}\cdots\boldsymbol{B}^{r_m}\boldsymbol{A}$;
2. $\boldsymbol{R}=\boldsymbol{A}\boldsymbol{B}^{r_1}\boldsymbol{A}\boldsymbol{B}^{r_2}\cdots\boldsymbol{A}\boldsymbol{B}^{r_m}$;
3. $\boldsymbol{R}=\boldsymbol{B}^{r_1}\boldsymbol{A}\boldsymbol{B}^{r_2}\boldsymbol{A}\cdots\boldsymbol{B}^{r_{m-1}}\boldsymbol{A}\boldsymbol{B}^{r_m}$;
4. $\boldsymbol{R}=\boldsymbol{A}\boldsymbol{B}^{r_1}\boldsymbol{A}\boldsymbol{B}^{r_2}\cdots\boldsymbol{A}\boldsymbol{B}^{r_m}\boldsymbol{A}$.

where $r_1,\dots,r_m\in\{1,2\}$. This gives the desired conclusion. We shall call any one of these four representations a ***reduced representation***. It is still possible, after reduction to a product in $\{\boldsymbol{B}\boldsymbol{A},\boldsymbol{B}^2\boldsymbol{A}\}$, that the representation as such a product will not be unique. For example, of $\theta=\pi$ then we have $\boldsymbol{B}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{B}^2$. The following result gives a condition under which this lack of uniqueness cannot happen.

5.5.38 Lemma *If* $\cos\theta$ *is transcendental, i.e., it is not a root of any polynomial with rational coefficients, then for* $\mathbf{R} \in \mathsf{G} \setminus \{\mathbf{I}_3\}$ *there exists a unique reduced representation*

$$\mathbf{R} = \mathbf{R}_1 \cdots \mathbf{R}_\mathrm{k}$$

for $\mathrm{k} \in \mathbb{Z}_{>0}$ *and* $\mathbf{R}_\mathrm{j} \in \{\mathbf{A}, \mathbf{B}, \mathbf{B}^2\}$, $\mathrm{j} \in \{1, \dots, \mathrm{k}\}$.

Proof The existence of the representation follows from the fact that $\boldsymbol{A}^{-1} = \boldsymbol{A}$ and $\boldsymbol{B}^{-1} = \boldsymbol{B}^2$. Thus we need only show uniqueness. It suffices to show that it is not possible to write

$$\boldsymbol{I}_3 = \boldsymbol{R}_1 \cdots \boldsymbol{R}_k$$

for $k \in \mathbb{Z}_{>0}$ and $\boldsymbol{R}_j \in \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}^2\}$, $j \in \{1, \dots, k\}$. Indeed, if we have

$$\boldsymbol{R}_1 \cdots \boldsymbol{R}_k = \boldsymbol{R}'_1 \cdots \boldsymbol{R}'_{k'}$$

with the factors in the products not being identical on the left and right, then

$$\boldsymbol{I}_3 = \boldsymbol{R}_k^{-1} \cdots \boldsymbol{R}_1^{-1} \boldsymbol{R}'_1 \cdots \boldsymbol{R}'_{k'},$$

giving $\boldsymbol{I}_3$ as a product in the factors $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}^2\}$.

It is clear that $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}^2 \neq \boldsymbol{I}_3$.

Now let $\boldsymbol{R}$ be one of the first of the four reduced representations given preceding the statement of the lemma. Thus $\boldsymbol{R} = \boldsymbol{R}_1 \cdots \boldsymbol{R}_k$ with $\boldsymbol{R}_j \in \{\boldsymbol{B}\boldsymbol{A}, \boldsymbol{B}^2\boldsymbol{A}\}$, $j \in \{1, \dots, k\}$. By an elementary inductive computation on k one can check that the third component of the vector $\boldsymbol{R}\boldsymbol{e}_3$ is a polynomial in $\cos\theta$ whose coefficients are rational. Since $\cos\theta$ is transcendental is cannot hold that $\boldsymbol{R}\boldsymbol{e}_3 = \boldsymbol{e}_3$ and so $\boldsymbol{R} \neq \boldsymbol{I}_3$.

If $\boldsymbol{R}$ has the second of the four reduced representations then $\boldsymbol{R}' = \boldsymbol{A}\boldsymbol{R}\boldsymbol{A}$ has the first of the four forms, and so cannot be equal to $\boldsymbol{I}_3$. Therefore, $\boldsymbol{R} \neq \boldsymbol{I}_3$ since, if it did, we would have $\boldsymbol{R}' = \boldsymbol{A}^2 = \boldsymbol{I}_3$.

Next let $\boldsymbol{R}$ be of the third of the reduced representations, assuming that m has been chosen to be the smallest positive integer for which the representation is possible; note that we must have $m > 1$. Suppose that $\boldsymbol{R} = \boldsymbol{I}_3$. Note that

$$\boldsymbol{I}_3 = \boldsymbol{B}^{-r_1}\boldsymbol{R}\boldsymbol{B}^{r_1} = \boldsymbol{A}\boldsymbol{B}^{r_2} \cdots \boldsymbol{A}\boldsymbol{B}^{r_1+r_m}.$$

If $r_1 = r_m$ then $\boldsymbol{B}^{r_1+r_m} \in \{\boldsymbol{B}^2, \boldsymbol{B}^4 = \boldsymbol{B}\}$ and so this gives $\boldsymbol{I}_3$ as a reduced representation in the second of the four forms. This cannot be, so we cannot have $r_1 = r_m$. Therefore, the only other possibility is $r_1 + r_m = 3$. In this case, if $m > 3$ we have

$$\boldsymbol{I}_3 = \boldsymbol{A}\boldsymbol{B}^{r_m}\boldsymbol{R}\boldsymbol{B}^{r_1}\boldsymbol{A} = \boldsymbol{B}^{r_2}\boldsymbol{A} \cdots \boldsymbol{A}\boldsymbol{B}^{r_{m-1}},$$

contradicting our assumption that m is the smallest positive integer giving the reduced representation. Thus we must have $m \in \{2, 3\}$. For $m = 2$ we then have

$$\boldsymbol{I}_3 = \boldsymbol{B}^{r_2}\boldsymbol{R}\boldsymbol{B}^{r_1} = \boldsymbol{B}$$

and if $m = 3$ we then have

$$\boldsymbol{I}_3 = \boldsymbol{A}\boldsymbol{B}^{r_3}\boldsymbol{R}\boldsymbol{B}^{r_1}\boldsymbol{A} = \boldsymbol{B}^{r_2}.$$

Both of these conclusions are not possible, and so we cannot have $\boldsymbol{R} = \boldsymbol{I}_3$.

Finally, we consider $\boldsymbol{R}$ to have the fourth of the four reduced representations. In this case, if $\boldsymbol{R} = \boldsymbol{I}_3$ then $\boldsymbol{I}_3 = \boldsymbol{A}\boldsymbol{R}\boldsymbol{A}$ has the third reduced representation, giving a contradiction. ■

We now fix θ such that $\cos\theta$ is transcendental; this is possible since only a countable subset of numbers are not transcendental and since image(cos) = $[-1, 1]$. If $\boldsymbol{R} \in \mathsf{G}$, by the preceding lemma we can write

$$\boldsymbol{R} = \boldsymbol{R}_1 \cdots \boldsymbol{R}_k$$

for $\boldsymbol{R}_1, \dots, \boldsymbol{R}_k \in \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{B}^2\}$ with this representation being unique when it is reduced. In this case we call k the ***length*** of $\boldsymbol{R}$ which we denote by $\ell(\boldsymbol{R})$. The following lemma now uses the preceding lemma to give an essential decomposition of G.

5.5.39 Lemma *The group* G *has a partition* $(\mathsf{G}_1, \mathsf{G}_2, \mathsf{G}_3)$ *into three nonempty subsets such that*

(i) $\mathbf{R} \in \mathsf{G}_1$ *if and only if* $\mathbf{AR} \in \mathsf{G}_2 \cup \mathsf{G}_3$*;*

(ii) $\mathbf{R} \in \mathsf{G}_1$ *if and only if* $\mathbf{BR} \in \mathsf{G}_2$*;*

(iii) $\mathbf{R} \in \mathsf{G}_1$ *if and only if* $\mathbf{B}^2\mathbf{R} \in \mathsf{G}_3$*.*

Proof We define the partitions inductively by the length of their elements. For $\ell(\boldsymbol{R}) = 1$ we assign

$$\boldsymbol{I}_3 \in \mathsf{G}_1,\ \boldsymbol{A} \in \mathsf{G}_2,\ \boldsymbol{B} \in \mathsf{G}_2,\ \boldsymbol{B}^2 \in \mathsf{G}_3. \tag{5.9}$$

Now suppose that all elements $\boldsymbol{R} \in \mathsf{G}$ for which $\ell(\boldsymbol{R}) = m$ have been assigned to G_1, G_2, or G_3. If $\ell(\boldsymbol{R}) = m+1$ then write the reduced representation of $\boldsymbol{R}$ as $\boldsymbol{R} = \boldsymbol{R}_1 \cdots \boldsymbol{R}_{m+1}$. Let $\boldsymbol{R}' = \boldsymbol{R}_2 \cdots \boldsymbol{R}_{m+1}$ so that $\ell(\boldsymbol{R}') = m$. We then assign $\boldsymbol{R}$ to either G_1, G_2, or G_3 as follows:

$$\begin{aligned}
\boldsymbol{R}_1 = \boldsymbol{A},\ \boldsymbol{R}_2 \in \{\boldsymbol{B}, \boldsymbol{B}^2\},\ \boldsymbol{R}' \in \mathsf{G}_1 \quad &\Longrightarrow \quad \boldsymbol{R} \in \mathsf{G}_2,\\
\boldsymbol{R}_1 = \boldsymbol{A},\ \boldsymbol{R}_2 \in \{\boldsymbol{B}, \boldsymbol{B}^2\},\ \boldsymbol{R}' \in \mathsf{G}_2 \cup \mathsf{G}_3 \quad &\Longrightarrow \quad \boldsymbol{R} \in \mathsf{G}_1, && (5.10)\\
\boldsymbol{R}_1 = \boldsymbol{B},\ \boldsymbol{R}_2 = \boldsymbol{A},\ \boldsymbol{R}' \in \mathsf{G}_j, \quad &\Longrightarrow \quad \boldsymbol{R} \in \mathsf{G}_{j+1}, && (5.11)\\
\boldsymbol{R}_1 = \boldsymbol{B}^2,\ \boldsymbol{R}_2 = \boldsymbol{A},\ \boldsymbol{R}' \in \mathsf{G}_j, \quad &\Longrightarrow \quad \boldsymbol{R} \in \mathsf{G}_{j+2}, && (5.12)
\end{aligned}$$

where we adopt the notational convention that $\mathsf{G}_4 = \mathsf{G}_1$ and $\mathsf{G}_5 = \mathsf{G}_2$. Doing this for each m gives subsets G_1, G_2, and G_3 of G whose union equals G. Moreover, one can check that our inductive construction is unambiguous and so assigns each $\boldsymbol{R} \in \mathsf{G}$ to a unique component G_1, G_2, or G_3. It remains to show that the partition defined has the desired properties (i)–(iii).

We do this by induction on the length of the elements of G. It is obviously true for elements of length 1, using the rules prescribed above for forming the partitions. Now suppose that if $\boldsymbol{R} \in \mathsf{G}$ has length less than $m \in \mathbb{Z}_{>0}$ we have verified properties (i)–(iii). We then let $\boldsymbol{R} \in \mathsf{G}$ with $\boldsymbol{R} = \boldsymbol{R}_1 \cdots \boldsymbol{R}_m$ the unique reduced representation. We denote $\boldsymbol{R}' = \boldsymbol{R}_2 \cdots \boldsymbol{R}_m$. We consider various cases.

1. $\boldsymbol{R}_1 = \boldsymbol{A}$: We have $\boldsymbol{AR} = \boldsymbol{R}'$. Thus $\ell(\boldsymbol{AR}) = m - 1$ and so the induction hypothesis can be applied to $\boldsymbol{AR}$. Doing so yields

$$\begin{aligned}
\boldsymbol{R} \notin \mathsf{G}_1 \quad \Longleftrightarrow \quad \boldsymbol{A}^2\boldsymbol{R} \notin \mathsf{G}_1 \quad &\Longleftrightarrow \quad \boldsymbol{A}(\boldsymbol{AR}) \in \mathsf{G}_2 \cup \mathsf{G}_3\\
\Longleftrightarrow \quad \boldsymbol{AR} \in \mathsf{G}_1 \quad &\Longleftrightarrow \quad \boldsymbol{AR} \notin \mathsf{G}_2 \cup \mathsf{G}_3.
\end{aligned}$$

Thus $\boldsymbol{R} \in \mathsf{G}_1$ if and only if $\boldsymbol{AR} \in \mathsf{G}_2 \cup \mathsf{G}_3$ and so (i) holds. Moreover, (5.11) and (5.12) give

$$\boldsymbol{BR} \in \mathsf{G}_2 \quad \Longleftrightarrow \quad \boldsymbol{R} \in \mathsf{G}_1, \quad \boldsymbol{B}^2\boldsymbol{R} \in \mathsf{G}_3 \quad \Longleftrightarrow \quad \boldsymbol{R} \in \mathsf{G}_1,$$

which gives properties (ii) and (iii).

2. $\boldsymbol{R}_1 = \boldsymbol{B}$: In this case, (5.10) immediately gives

$$\boldsymbol{A}\boldsymbol{R} \in \mathsf{G}_2 \cup \mathsf{G}_3 \iff \boldsymbol{R} \in \mathsf{G}_1,$$

which gives condition (i). We also have $\boldsymbol{B}\boldsymbol{R} = \boldsymbol{B}^2\boldsymbol{R}'$ with $\ell(\boldsymbol{R}') = m - 1$ and with $\boldsymbol{R}_2 = \boldsymbol{A}$. Thus we can apply (5.11) and (5.12) to get

$$\begin{aligned} \boldsymbol{B}\boldsymbol{R} \in \mathsf{G}_2 &\iff \boldsymbol{B}^2\boldsymbol{R}' \in \mathsf{G}_2 \iff \boldsymbol{B}^2\boldsymbol{R}' \in \mathsf{G}_5 \\ &\iff \boldsymbol{R}' \in \mathsf{G}_3, \iff \boldsymbol{B}\boldsymbol{R}' \in \mathsf{G}_4 \\ &\iff \boldsymbol{B}\boldsymbol{R}' \in \mathsf{G}_1 \iff \boldsymbol{R} \in \mathsf{G}_1 \end{aligned}$$

which gives condition (ii). We also immediately have, borrowing an implication from the preceding line,

$$\boldsymbol{B}^2\boldsymbol{R} \in \mathsf{G}_3 \iff \boldsymbol{R}' \in \mathsf{G}_3 \iff \boldsymbol{R} \in \mathsf{G}_1$$

giving condition (iii).

3. $\boldsymbol{R}_1 = \boldsymbol{B}^2$: From (5.10) we have

$$\boldsymbol{A}\boldsymbol{R} \in \mathsf{G}_2 \cup \mathsf{G}_3 \iff \boldsymbol{R} \in \mathsf{G}_1,$$

which gives condition (i). We have $\boldsymbol{B}\boldsymbol{R} = \boldsymbol{R}'$ with $\ell(\boldsymbol{R}') = m - 1$ and with $\boldsymbol{R}_2 = \boldsymbol{A}$. We then have, using (5.11) and (5.12),

$$\begin{aligned} \boldsymbol{B}\boldsymbol{R} \in \mathsf{G}_2 &\iff \boldsymbol{R}' \in \mathsf{G}_2 \iff \boldsymbol{B}^2\boldsymbol{R}' \in \mathsf{G}_4 \\ &\iff \boldsymbol{B}^2\boldsymbol{R}' \in \mathsf{G}_1 \iff \boldsymbol{R} \in \mathsf{G}_1 \end{aligned}$$

which gives condition (ii). Finally, we have

$$\boldsymbol{B}^2\boldsymbol{R} \in \mathsf{G}_3 \iff \boldsymbol{B}\boldsymbol{R}' \in \mathsf{G}_3 \iff \boldsymbol{R}' \in \mathsf{G}_2 \iff \boldsymbol{R} \in \mathsf{G}_1,$$

borrowing an implication from the preceding line. Thus we also have condition (iii). ■

Now we state the result on which the entire proof hinges. It relates the algebraic constructions thus far seen in the proof to conclusions about subsets of $\mathbb{S}^2$. It is here that we employ the Axiom of Choice in an essential way.

5.5.40 Lemma *There exists a partition* $(\mathrm{P}, \mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3)$ *of* $\mathbb{S}^2$ *for which*

(i) P *is countable,*

(ii) $\mathbf{A}(\mathrm{S}_1) = \mathrm{S}_2 \cup \mathrm{S}_3$*,*

(iii) $\mathbf{B}(\mathrm{S}_1) = \mathrm{S}_2$*, and*

(iv) $\mathbf{B}^2(\mathrm{S}_2) = \mathrm{S}_3$.

Proof Define

$$P = \{\boldsymbol{x} \in \mathbb{S}^2 \mid \boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{x},\ \boldsymbol{R} \in \mathsf{G} \setminus \{\boldsymbol{I}_3\}\}.$$

Since G is countable and since $\boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{x}$ for two point $\boldsymbol{x} \in \mathbb{S}^2$ by Exercise **??**, it follows that P is countable. If $\boldsymbol{x} \in \mathbb{S}^2 \setminus P$ denote

$$\mathsf{G}\boldsymbol{x} = \{\boldsymbol{R}(\boldsymbol{x}) \mid \boldsymbol{R} \in \mathsf{G}\}.$$

We claim that $\mathsf{G}\boldsymbol{x} \subseteq \mathbb{S}^2 \setminus P$. Indeed, suppose otherwise. Then there exists $\boldsymbol{R} \in \mathsf{G}$ such that $\boldsymbol{R}(\boldsymbol{x}) \in P$. Then $\boldsymbol{SR}(\boldsymbol{x}) = \boldsymbol{R}(\boldsymbol{x})$ for $\boldsymbol{S} \in \mathsf{G} \setminus \{\boldsymbol{I}_3\}$. Then $\boldsymbol{R}^{-1}\boldsymbol{SR}(\boldsymbol{x}) = \boldsymbol{x}$ with $\boldsymbol{R}^{-1}\boldsymbol{SR} \neq \boldsymbol{I}_3$, contradicting the assumption that $\boldsymbol{x} \notin P$. If $\boldsymbol{x}, \boldsymbol{y} \in P$, we claim that either $\mathsf{G}\boldsymbol{x} = \mathsf{G}\boldsymbol{y}$ or $\mathsf{G}\boldsymbol{x} \cap \mathsf{G}\boldsymbol{y} = \emptyset$. Indeed, suppose that $\boldsymbol{z} \in \mathsf{G}\boldsymbol{x} \cap \mathsf{G}\boldsymbol{y}$ so that $\boldsymbol{z} = \boldsymbol{Rx} = \boldsymbol{Sy}$ for $\boldsymbol{R}, \boldsymbol{S} \in \mathsf{G}$. Then let $\boldsymbol{z} \in \mathsf{G}\boldsymbol{x}$ with $\boldsymbol{z} = \boldsymbol{Tx}$. Then $\boldsymbol{z} = \boldsymbol{TR}^{-1}\boldsymbol{Sy}$ and so $\boldsymbol{z} \in \mathsf{G}\boldsymbol{y}$. Thus

$$\{\mathsf{G}\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{S}^2 \setminus P\} \tag{5.13}$$

is a partition of $\mathbb{S}^2 \setminus P$. Let $C \subseteq \mathbb{S}^2 \setminus P$ be chosen so that it contains exactly one element of each component of this partition, using the Axiom of Choice. Now define

$$S_j = \{\boldsymbol{Rx} \mid \boldsymbol{x} \in C,\ \boldsymbol{R} \in \mathsf{G}_j\}, \qquad j \in \{1,2,3\}.$$

We claim that $\mathbb{S}^2 \setminus P = S_1 \cup S_2 \cup S_3$. Indeed, let $\boldsymbol{x} \in \mathbb{S}^2 \setminus P$. Then $\boldsymbol{x} = \boldsymbol{R}(\boldsymbol{x}')$ for some $\boldsymbol{x}' \in C$ and for some $\boldsymbol{R} \in \mathsf{G}$. Since $\mathsf{G} = \mathsf{G}_1 \cup \mathsf{G}_2 \cup \mathsf{G}_3$ it follows that $\boldsymbol{x} \in S_j$ for some $j \in \{1,2,3\}$. We also claim that $S_j \cap S_k = \emptyset$ for $j \neq k$. Indeed, suppose that $\boldsymbol{x} \in S_j \cap S_k$. Then $\boldsymbol{x} = \boldsymbol{R}_j(\boldsymbol{x}_j) = \boldsymbol{R}_k(\boldsymbol{x}_k)$ for some $\boldsymbol{R}_j \in \mathsf{G}_j$, $\boldsymbol{R}_k \in \mathsf{G}_k$, $\boldsymbol{x}_j, \boldsymbol{x}_k \in C$. Since C contains exactly one element from each component in the partition (5.13), it follows from the fact that $\boldsymbol{x}_j = \boldsymbol{R}_j^{-1}\boldsymbol{R}_k\boldsymbol{x}_k$ that $\boldsymbol{x}_j$ and $\boldsymbol{x}_k$ are in the same component of the partition and so are equal. Since $C \subseteq \mathbb{S}^2 \setminus P$ it follows that $\boldsymbol{R}_j^{-1}\boldsymbol{R}_k = \boldsymbol{I}_3$ and so $\boldsymbol{R}_j = \boldsymbol{R}_k$. Thus $j = k$. This shows that (P, S_1, S_2, S_3) is indeed a partition of $\mathbb{S}^2$.

Moreover, we compute

$$\begin{aligned}
\boldsymbol{A}(S_1) &= \{\boldsymbol{AR}(\boldsymbol{x}) \mid \boldsymbol{R} \in \mathsf{G}_1,\ \boldsymbol{x} \in C\} = \{\boldsymbol{Tx} \mid \boldsymbol{T} \in \mathsf{G}_2 \cup \mathsf{G}_3,\ \boldsymbol{x} \in C\} = S_2 \cup S_3,\\
\boldsymbol{B}(S_1) &= \{\boldsymbol{BR}(\boldsymbol{x}) \mid \boldsymbol{R} \in \mathsf{G}_1,\ \boldsymbol{x} \in C\} = \{\boldsymbol{Tx} \mid \boldsymbol{T} \in \mathsf{G}_2,\ \boldsymbol{x} \in C\} = S_2,\\
\boldsymbol{B}^2(S_2) &= \{\boldsymbol{B}^2\boldsymbol{R}(\boldsymbol{x}) \mid \boldsymbol{R} \in \mathsf{G}_1,\ \boldsymbol{x} \in C\} = \{\boldsymbol{Tx} \mid \boldsymbol{T} \in \mathsf{G}_3,\ \boldsymbol{x} \in C\} = S_3,
\end{aligned}$$

giving conditions (ii)–(iv). ■

The following rather technical lemma will be crucial to our proof.

5.5.41 Lemma *If* $\mathrm{P} \subseteq \mathbb{S}^2$ *is countable then there exists* $\mathrm{Q} \subseteq \mathbb{S}^2$ *countable and* $\mathbf{T} \in \mathsf{O}(3)$ *such that* $\mathrm{P} \subseteq \mathrm{Q}$ *and such that* $\mathbf{T}(\mathrm{Q}) = \mathrm{Q} - \mathrm{P}$.

Proof Let $\boldsymbol{v} = (v_1, v_2, 0) \in \mathbb{S}^1$ be such that $\boldsymbol{v}, -\boldsymbol{v} \notin P$; since P is countable this is possible. Define

$$\boldsymbol{T}_0 = \begin{bmatrix} v_1 & v_2 & 0 \\ -v_2 & v_1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and note that $\boldsymbol{T}_0 \in \mathsf{O}(3)$ since $v_1^2 + v_2^2 = 1$. Note that $\boldsymbol{T}_0(\boldsymbol{v}) = \boldsymbol{e}_1$ and that $\boldsymbol{e}_1, -\boldsymbol{e}_1 \notin \boldsymbol{T}_0(P)$. For $t \in \mathbb{R}$ define the orthogonal matrix

$$\boldsymbol{U}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{bmatrix}.$$

Note that

$$\boldsymbol{U}_t^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(kt) & -\sin(kt) \\ 0 & \sin(kt) & \cos(kt) \end{bmatrix}, \qquad k \in \mathbb{Z}_{>0}.$$

For $\boldsymbol{x}, \boldsymbol{y} \in P$ and for $k \in \mathbb{Z}_{>0}$ consider the equation $\boldsymbol{U}_t^k(\boldsymbol{x}) = \boldsymbol{y}$. In components this equation reads

$$x_1 = y_1,\ \cos(kt)x_2 - \sin(kt)x_3 = y_2,\ \sin(kt)x_2 + \cos(kt)x_3 = y_3.$$

If $y_1 \neq x_1$ then these equations have no solution in t. If $y_1 = x_1$ then there are infinitely many solutions in t, all satisfying $\cos(kt) = 1$ and $\sin(kt) = 0$. In particular, in $[0, 2\pi)$ there are exactly k solutions if $y_1 = x_1$. Therefore, since the set $P \times P \times \mathbb{Z}_{>0}$ is countable by Proposition **??**, it follows that the complement to the set

$$\left\{t \in \mathbb{R} \;\middle|\; \boldsymbol{T}_0(P) \cap \left(\cup_{k \in \mathbb{Z}_{>0}} \boldsymbol{U}_t^k(\boldsymbol{T}_0(P))\right) = \emptyset\right\} \tag{5.14}$$

is countable. Thus choose a t in the set (5.14) and denote $\boldsymbol{U} = \boldsymbol{U}_t$. Then define $\boldsymbol{T} = \boldsymbol{T}_0^{-1}\boldsymbol{U}\boldsymbol{T}_0$ and

$$Q = P \cup \left(\cup_{k \in \mathbb{Z}_{>0}} \boldsymbol{T}^k(P)\right).$$

One can directly check that $\boldsymbol{U}^k\boldsymbol{T}_0 = \boldsymbol{T}_0\boldsymbol{T}^k$ for $k \in \mathbb{Z}_{>0}$. Thus, using the fact that $\boldsymbol{U}$ is defined by t satisfying (5.14), we have

$$\boldsymbol{T}_0(P \cap \boldsymbol{T}(Q)) = \boldsymbol{T}_0\Big(P \cap \left(\cup_{k \in \mathbb{Z}_{>0}} \boldsymbol{T}^k(P)\right)\Big) = \emptyset.$$

We then conclude that $P \cap \boldsymbol{T}(Q) = \emptyset$, and since $Q = P \cup \boldsymbol{T}(Q)$, as follows from the definition of Q, it follows that $\boldsymbol{T}(Q) = Q - P$. ■

Using the previous lemma, we now make a decomposition of $\mathbb{S}^2$.

5.5.42 Lemma *There exists a partition* $(\mathrm{T_j})_{1\leq j\leq 10}$ *of* $\mathbb{S}^2$ *and isometries* $\boldsymbol{\sigma}_1, \dots, \boldsymbol{\sigma}_{10} \in \mathsf{E}(\mathrm{n})$ *such that* $(\boldsymbol{\sigma}_\mathrm{j}(\mathrm{T_j}))_{1\leq j\leq 6}$ *and* $(\boldsymbol{\sigma}_\mathrm{j}(\mathrm{T_j}))_{7\leq j\leq 10}$ *are both partitions of* $\mathbb{S}^2$.

Proof We use the partition (P, S_1, S_2, S_3) from Lemma 5.5.40. We define

$$U_1 = \boldsymbol{A}(S_2),\ U_2 = \boldsymbol{B}\boldsymbol{A}(S_2),\ U_3 = \boldsymbol{B}^2\boldsymbol{A}(S_2),$$
$$V_1 = \boldsymbol{A}(S_3),\ U_2 = \boldsymbol{B}\boldsymbol{A}(S_3),\ U_3 = \boldsymbol{B}^2\boldsymbol{A}(S_3).$$

By Lemma 5.5.40 we see that (U_j, V_j) is a partition of S_j for each $j \in \{1, 2, 3\}$. Now define

$$T_7 = U_1,\ T_8 = U_2,\ T_9 = U_3,\ T_{10} = P,$$
$$\boldsymbol{\sigma}_7 = \boldsymbol{B}^2\boldsymbol{A},\ \boldsymbol{\sigma}_8 = \boldsymbol{A}\boldsymbol{B}^2,\ \boldsymbol{\rho}_9 = \boldsymbol{B}\boldsymbol{A}\boldsymbol{B},\ \boldsymbol{\rho}_{10} = \boldsymbol{I}_3.$$

We can then check that $\boldsymbol{\sigma}_{10}(T_{10}) = P$ and $\boldsymbol{\sigma}_j(T_j) = S_{j-6}$ for $j \in \{7, 8, 9\}$. Thus $(\boldsymbol{\sigma}_j(T_j))_{7\leq j\leq 10}$ is a partition of $\mathbb{S}^2$. Next note that

$$\mathbb{S}^2 \setminus (T_7 \cup T_8 \cup T_9 \cup T_{10}) = V_1 \cup V_2 \cup V_3.$$

Let $Q \subseteq \mathbb{S}^2$ and $\boldsymbol{T} \in \mathsf{O}(3)$ be as in Lemma 5.5.41 and define

$$T_1 = \boldsymbol{\sigma}_8(S_1 \cap Q),\ T_2 = \boldsymbol{\sigma}_9(S_2 \cap Q),\ T_3 = \boldsymbol{\sigma}_7(S_3 \cap Q),$$
$$T_1 = \boldsymbol{\sigma}_8(S_1 \setminus Q),\ T_2 = \boldsymbol{\sigma}_9(S_2 \setminus Q),\ T_3 = \boldsymbol{\sigma}_7(S_3 \setminus Q).$$

Then (T_1, T_4) partitions $\boldsymbol{\rho}_8(S_1) = V_1$, (T_2, T_5) partitions $\boldsymbol{\rho}_9(S_2) = V_2$, and (T_3, T_6) partitions $\boldsymbol{\rho}_7(S_3) = V_3$. Therefore, $(T_j)_{1\le j\le 10}$ is a partition of $\mathbb{S}^2$. Finally, define

$$\boldsymbol{\sigma}_4 = \boldsymbol{\sigma}_8^{-1},\ \boldsymbol{\sigma}_5 = \boldsymbol{\sigma}_9^{-1},\ \boldsymbol{\sigma}_6 = \boldsymbol{\sigma}_7^{-1},\ \boldsymbol{\sigma}_j = \boldsymbol{R}\boldsymbol{\sigma}_{j+3}, \qquad j \in \{1,2,3\}.$$

One can then directly check that $\boldsymbol{\sigma}_{j+3}(T_{j+3}) = S_j \setminus Q$, $j \in \{1,2,3\}$ so that

$$\cup_{j=1}^3 \boldsymbol{\sigma}_{j+3}(T_{j+3}) = \mathbb{S}^2 \setminus Q$$

by virtue of the fact that $P \subseteq Q$. Moreover,

$$\boldsymbol{\sigma}_j(T_j) = \boldsymbol{R}^{-1}\boldsymbol{\sigma}_{j+3}(T_j) = \boldsymbol{R}^{-1}(S_j \cap Q), \qquad j \in \{1,2,3\},$$

which shows that $\boldsymbol{\sigma}_j(T_j) \cap \boldsymbol{\sigma}_k(T_k) = \emptyset$ if $j \neq k$. This also shows that

$$\cup_{j=1}^3 \boldsymbol{\sigma}_j(T_j) = \boldsymbol{R}^{-1}(Q - P) = Q.$$

Thus $(\boldsymbol{\sigma}_j(T_j))_{1\le j\le 6}$ is a partition of $\mathbb{S}^2$, completing the proof. ■

5.5.43 Lemma *For* $\mathrm{r} \in \mathbb{R}_{>0}$ *and* $\mathbf{x}_0 \in \mathbb{R}^3$ *there exists a partition* $(\mathrm{B_j})_{1\le j\le 40}$ *of* $\overline{\mathsf{B}}^3(\mathrm{r}, \mathbf{x}_0)$ *and isometries* $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_{40} \in \mathsf{E}(\mathrm{n})$ *such that* $(\boldsymbol{\rho}_\mathrm{j}(\mathrm{B_j}))_{1\le j\le 24}$ *and* $(\boldsymbol{\rho}_\mathrm{j}(\mathrm{B_j}))_{25\le j\le 40}$ *are both partitions of* $\overline{\mathsf{B}}^3(\mathrm{r}, \mathbf{x}_0)$.

Proof Let us first prove the result when $r = 1$ and $\boldsymbol{x}_0 = \boldsymbol{0}$ in which case $\mathrm{bd}(\overline{\mathsf{B}}^3(1, \boldsymbol{0})) = \mathbb{S}^2$. For $S \subseteq \mathbb{S}^2$ let us denote

$$\hat{S} = \{\lambda \boldsymbol{x} \mid \lambda \in (0,1],\ \boldsymbol{x} \in S\}.$$

Thus, for example, $\hat{\mathbb{S}}^2 = \overline{\mathsf{B}}^3(1, \boldsymbol{0}) \setminus \{\boldsymbol{0}\}$.

Let $P = \{\boldsymbol{e}_1\}$ and, by Lemma 5.5.41, let $Q \subseteq \mathbb{S}^2$ and $\boldsymbol{R}_0 \in \mathsf{O}(3)$ be such that Q is countable, $P \subseteq Q$, and $\boldsymbol{R}_0(Q) = Q \setminus P$. Define

$$N_1 = \left\{\tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{e}_1) \,\middle|\, \boldsymbol{x} \in Q\right\}$$

and define $\boldsymbol{\rho}_0 \in \mathsf{E}(3)$ by

$$\boldsymbol{\rho}_0(\boldsymbol{x}) = \boldsymbol{R}_0(\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{e}_1) - \tfrac{1}{2}\boldsymbol{e}_1;$$

thus $\boldsymbol{\rho}_0$ is a rotation about $\frac{1}{2}\boldsymbol{e}_1$. Note that $\boldsymbol{0} \in N_1$ and that $\boldsymbol{\rho}_0(N_1) = N_1 \setminus \{\boldsymbol{0}\}$. Denote

$$N_2 = \overline{\mathsf{B}}^3(1, \boldsymbol{0}) \setminus N_1,\ \boldsymbol{f}_1 = \boldsymbol{\rho}_0,\ \boldsymbol{f}_2 = \boldsymbol{I}_3,\ M_k = \boldsymbol{f}_k(N_k), \qquad j \in \{1,2\}.$$

Then we have (N_1, N_2) as a partition of $\overline{\mathsf{B}}^3(1, \boldsymbol{0})$ and (M_1, M_2) as a partition of $\hat{\mathbb{S}}^2$. Let $(T_j)_{1\le j\le 10}$ and $(\boldsymbol{\sigma}_j)_{1\le j\le 10}$ be as in Lemma 5.5.42, noting that $(\hat{T}_j)_{1\le j\le 10}$ is a partition of $\hat{\mathbb{S}}^2$.

Note that

$$(M_k \cap \hat{T}_j \cap \boldsymbol{\sigma}_j(M_l) \mid\ k, l \in \{1,2\},\ j \in \{1, \dots, 10\})$$

is a partition of $\hat{\mathbb{S}}^2$ into forty components. Moreover, if we define

$$B_{klj} = \boldsymbol{f}_k^{-1}(M_k \cap \hat{T}_j \cap \boldsymbol{\sigma}_j^{-1}(M_l)), \qquad k, l \in \{1,2\},\ j \in \{1, \dots, 10\},$$

then these sets partition $\overline{\mathsf{B}}^3(1,\mathbf{0})$. Moreover, for fixed $j \in \{1,\dots,10\}$, the sets

$$\boldsymbol{\sigma}_j \circ \boldsymbol{f}_k(B_{klj}) = M_l \cap \boldsymbol{\sigma}_j(M_k \cap \hat{T}_j), \qquad k,l \in \{1,2\},$$

partition $\boldsymbol{\sigma}_j(\hat{T}_j)$. By Lemma 5.5.42 we have that

$$(\boldsymbol{\sigma}_j \circ \boldsymbol{f}_k(B_{klj}) \mid \ k,l \in \{1,2\},\ j \in \{1,\dots,6\}),$$
$$(\boldsymbol{\sigma}_j \circ \boldsymbol{f}_k(B_{klj}) \mid \ k,l \in \{1,2\},\ j \in \{7,\dots,10\})$$

each partition $\hat{\mathbb{S}}^2$. Therefore, if we define $\boldsymbol{\rho}_{klj} = \boldsymbol{f}_l^{-1} \circ \boldsymbol{\sigma}_j \circ \boldsymbol{f}_k$ we see that

$$(\boldsymbol{\rho}_{klj}(B_{klj}) \mid \ k,l \in \{1,2\},\ j \in \{1,\dots,6\}),$$
$$(\boldsymbol{\rho}_{klj}(B_{klj}) \mid \ k,l \in \{1,2\},\ j \in \{7,\dots,10\})$$

each partition $\overline{\mathsf{B}}^3(1,\mathbf{0})$. This proves the lemma for $r = 1$ and $\boldsymbol{x}_0 = \mathbf{0}$.

In general, we define $B'_{klj} \subseteq \overline{\mathsf{B}}^3(r,\boldsymbol{x}_0)$ and $\boldsymbol{\rho}'_{klj} \in \mathsf{E}(3)$, $k,l \in \{1,2\}$, $j \in \{1,\dots,10\}$, by

$$B'_{klj} = \{r\boldsymbol{x} + \boldsymbol{x}_0 \mid \ \boldsymbol{x} \in B_{klj}\}$$

and

$$\boldsymbol{\rho}'_{klj}(\boldsymbol{x}) = r\boldsymbol{\rho}_{klj}(r^{-1}(\boldsymbol{x} - \boldsymbol{x}_0)) + \boldsymbol{x}_0$$

and then directly verify that

$$(\boldsymbol{\rho}'_{klj}(B'_{klj}) \mid \ k,l \in \{1,2\},\ j \in \{1,\dots,6\}),$$
$$(\boldsymbol{\rho}'_{klj}(B'_{klj}) \mid \ k,l \in \{1,2\},\ j \in \{7,\dots,10\})$$

each partition $\overline{\mathsf{B}}^3(r,\boldsymbol{x}_0)$. ■

Now let us introduce some notation that will be convenient in the remainder of the proof. If set $X, Y \subseteq \mathbb{R}^n$ are piecewise congruent then we write $X \sim Y$. If X is piecewise congruent to a subset of Y then we write $X \precsim Y$. The following lemma records some useful facts about these relations.

5.5.44 Lemma *For* $X, Y, Z \subseteq \mathbb{R}^n$ *the following statements hold:*

(i) $X \sim X$;

(ii) *if* $X \sim Y$ *then* $Y \sim X$;

(iii) *if* $X \sim Y$ *and* $Y \sim Z$ *then* $X \sim Z$;

(iv) *if* $X \sim Y$ *then* $X \precsim Y$;

(v) *if* $X \precsim Y$ *and* $Y \precsim Z$ *then* $X \precsim Z$;

(vi) *if* $X \subseteq Y$ *then* $X \precsim Y$;

(vii) *if* $X \precsim Y$ *and* $Y \precsim X$ *then* $X \sim Y$.

Proof (i) This is obvious.

(ii) This follows since if $\boldsymbol{\rho}$ is an isometry then it is invertible and $\boldsymbol{\rho}^{-1}$ is an isometry.

(iii) Let $(X_1,\dots,X_N)$ and $(Y_1,\dots,Y_M)$ be partitions of X and Y, respectively, with $\boldsymbol{\rho}_1,\dots,\boldsymbol{\rho}_N \in \mathsf{E}(n)$ and $\boldsymbol{\sigma}_1,\dots,\sigma_M \in \mathsf{E}(n)$ such that $(\boldsymbol{\rho}_j(X_j))_{j\in\{1,\dots,N\}}$ and $(\boldsymbol{\sigma}_k(Y_k))_{k\in\{1,\dots,M\}}$ are partitions of Y and Z, respectively. Then, for $j \in \{1, \dots, N\}$ and $k \in \{1, \dots, M\}$, define $A_{jk} = X_j \cap \boldsymbol{\rho}_j^{-1}(Y_k)$, noting that the sets A_{jk}, $j \in \{1, \dots, N\}$, $k \in \{1, \dots, M\}$, form a partition of X. Thus the sets $\boldsymbol{\rho}_j(A_{jk}) = Y_k \cap \boldsymbol{\rho}_j(X_j)$, $j \in \{1, \dots, N\}$, $k \in \{1, \dots, M\}$, form a partition of Y and the sets $\boldsymbol{\sigma}_k \circ \boldsymbol{\rho}_j(A_{ij})$, $j \in \{1, \dots, N\}$, $k \in \{1, \dots, M\}$, form a partition of Z. Since $\boldsymbol{\sigma}_k \circ \rho_j \in \mathsf{E}(n)$ for each $j \in \{1, \dots, N\}$ and $k \in \{1, \dots, M\}$, it follows that $X \sim Z$, as desired.

(iv) This follows because $Y \subseteq Y$.

(v) Let $(X_1, \dots, X_N)$ and $(Y_1, \dots, Y_M)$ be partitions of X and Y, respectively, with $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_N \in \mathsf{E}(n)$ and $\boldsymbol{\sigma}_1, \dots, \sigma_M \in \mathsf{E}(n)$ such that $(\boldsymbol{\rho}_j(X_j))_{j \in \{1,\dots,N\}}$ and $(\boldsymbol{\sigma}_k(Y_k))_{k \in \{1,\dots,M\}}$ are partitions of $Y' \subseteq Y$ and $Z' \subseteq Z$, respectively. Then, for $j \in \{1, \dots, N\}$ and $k \in \{1, \dots, M\}$, define $A_{jk} = X_j \cap \boldsymbol{\rho}_j^{-1}(Y_k)$, noting that the sets A_{jk}, $j \in \{1, \dots, N\}$, $k \in \{1, \dots, M\}$, form a partition of X. Thus, for fixed $k \in \{1, \dots, M\}$, the sets $\boldsymbol{\rho}_j(A_{jk}) = Y_k \cap \boldsymbol{\rho}_j(X_j)$, $j \in \{1, \dots, N\}$, form a partition of $Y_k \cap Y'$ and the sets $\boldsymbol{\sigma}_k \circ \boldsymbol{\rho}_j(A_{ij})$, $j \in \{1, \dots, N\}$, $k \in \{1, \dots, M\}$, then form a partition for some subset $Z'' \subseteq Z$. Since $\boldsymbol{\sigma}_k \circ \rho_j \in \mathsf{E}(n)$ for each $j \in \{1, \dots, N\}$ and $k \in \{1, \dots, M\}$, it follows that $X \precsim Z$, as desired.

(vi) This is obvious.

(vii) Suppose that $X \sim Y'$ and $Y \sim X'$ for $X' \subseteq X$ and $Y' \subseteq Y$. Let $(X_1, \dots, X_N)$ and $(Y_1, \dots, Y_M)$ be partitions of X and Y, respectively, with $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_N \in \mathsf{E}(n)$ and $\boldsymbol{\sigma}_1, \dots, \sigma_M \in \mathsf{E}(n)$ such that $(\boldsymbol{\rho}_j(X_j))_{j \in \{1,\dots,N\}}$ and $(\boldsymbol{\sigma}_k(Y_k))_{k \in \{1,\dots,M\}}$ are partitions of $Y' \subseteq Y$ and $X' \subseteq X$, respectively. Define bijections $\boldsymbol{\rho}\colon X \to Y'$ and $\boldsymbol{\sigma}\colon Y \to X'$ by asking that $\boldsymbol{\rho}|X_j = \boldsymbol{\rho}_j$ and that $\boldsymbol{\sigma}|Y_k = \boldsymbol{\sigma}_k$ for $j \in \{1, \dots, N\}$ and $k \in \{1, \dots, M\}$. If $A \subseteq X$ denote

$$\tilde{A} = X \setminus \boldsymbol{\sigma}(Y \setminus \boldsymbol{\rho}(A)).$$

It is easy to verify that of $A \subseteq B \subseteq X$ that $\tilde{A} \subseteq \tilde{B}$. Now define $\mathscr{S} = \{A \subseteq X \mid A \subseteq \tilde{A}\}$. Since $\emptyset \in \mathscr{S}$, $\mathscr{S}$ is not empty. Let $S = \cup \mathscr{S}$. If $A \in \mathscr{S}$ then $A \subseteq S$ and so $\tilde{A} \subseteq \tilde{S}$. Therefore, $S \subseteq \tilde{S}$ and so $\tilde{S} \subseteq \tilde{\tilde{S}}$. Thus $\tilde{S} \in \mathscr{S}$ and so $\tilde{S} \subseteq S$. Therefore, $S = \tilde{S}$. Therefore, by definition of $\tilde{\cdot}$,

$$S = X \setminus \boldsymbol{\sigma}(Y \setminus \boldsymbol{\rho}(S)) \quad \Longrightarrow \quad X \setminus S = \boldsymbol{\sigma}(Y \setminus \boldsymbol{\rho}(S)).$$

This implies that $X \setminus S \in X'$. Now let $l \in \{1, \dots, N+M\}$ and define

$$A_l = \begin{cases} S \cap X_l, & l \in \{1, \dots, N\}, \\ \boldsymbol{\sigma}_{l-N}(Y_{l-N} \setminus \boldsymbol{\rho}(S)), & l \in \{N+1, \dots, N+M\} \end{cases}$$

and

$$\boldsymbol{\tau}_l = \begin{cases} \boldsymbol{\rho}_l, & l \in \{1, \dots, N\}, \\ \boldsymbol{\sigma}_{l-N}^{-1}, & l \in \{N+1, \dots, N+M\}. \end{cases}$$

One then verifies that $(A_1, \dots, A_N)$ partitions S, $(A_{N+1}, \dots, A_{N+M})$ partitions $X \setminus S$, $(\boldsymbol{\tau}_1(A_1), \dots, \boldsymbol{\tau}_N(A_N))$ partitions $\boldsymbol{\rho}(S)$, and $(\boldsymbol{\tau}_{N+1}(A_{N+1}), \dots, \boldsymbol{\tau}_{N+M}(A_{N+M}))$ partitions $Y \setminus \boldsymbol{\rho}(S)$. This gives $X \sim Y$. ■

Next we state a lemma about piecewise congruence of identical balls with finite unions of the same sized balls.

5.5.45 Lemma *If* $\mathrm{r} \in \mathbb{R}_{>0}$ *and* $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ *then* $\overline{\mathsf{B}}^3(\mathrm{r}, \mathbf{x}_0) \sim \cup_{j=1}^{k} \overline{\mathsf{B}}^3(\mathrm{r}, \mathbf{x}_j)$.

Proof We first prove the lemma in the case of $k = 2$, assuming that $\|x_1 - x_2\|_{\mathbb{R}^3} > 2\epsilon$, i.e., assuming that $\overline{\mathsf{B}}^3(r, x_1)$ and $\overline{\mathsf{B}}^3(r, x_2)$ do not intersect. Let $(B_j)_{1\le j\le 40}$ be the partition of $\overline{\mathsf{B}}^3(r, x_0)$ and let $(\rho_j)_{1\le j\le 40}$ be the isometries given by Lemma 5.5.43. Then define $\sigma_j \in \mathsf{E}(n)$, $j \in \{1, \dots, 40\}$, by

$$\sigma_j(x) = \begin{cases} \rho_j(x) - x_0 + x_1, & j \in \{1, \dots, 24\}, \\ \rho_j(x) - x_0 + x_2, & j \in \{25, \dots, 40\}. \end{cases}$$

Then $(\sigma_j(B_j))_{1\le j\le 24}$ is a partition of $\overline{\mathsf{B}}^3(r, x_1)$ and $(\sigma_j(B_j))_{25\le j\le 40}$ is a partition of $\overline{\mathsf{B}}^3(r, x_2)$ by Lemma 5.5.43. Thus we have $\overline{\mathsf{B}}^3(r, x_0) \sim \overline{\mathsf{B}}^3(r, x_1) \cup \overline{\mathsf{B}}^3(r, x_2)$ under the stated hypotheses.

Now we prove the lemma by induction on k. The result is clear for $k = 1$: one need only translate $\overline{\mathsf{B}}^3(r, x_0)$ to $\overline{\mathsf{B}}^3(r, x_1)$ by the isometry $x \mapsto x - x_0 + x_1$. Suppose the result holds for $k = m - 1$ and consider $x_1, \dots, x_m \in \mathbb{R}^n$. Choose an arbitrary $x_0' \in \mathbb{R}^n$ such that $\|x_0 - x_0'\|_{\mathbb{R}^3} > 2\epsilon$. By the induction hypothesis we have $\overline{\mathsf{B}}^3(r, x_0) \sim \cup_{j=1}^{m-1} \overline{\mathsf{B}}^3(r, x_j)$. Note that

$$\overline{\mathsf{B}}^3(r, x_m) \setminus \left(\cup_{j=1}^{m-1} \overline{\mathsf{B}}^3(r, x_j)\right) \subseteq \overline{\mathsf{B}}^3(r, x_m) \sim \overline{\mathsf{B}}^3(r, x_0'),$$

and so

$$\overline{\mathsf{B}}^3(r, x_m) \setminus \left(\cup_{j=1}^{m-1} \overline{\mathsf{B}}^3(r, x_j)\right) \precsim \overline{\mathsf{B}}^3(r, x_0').$$

From this we conclude that

$$\overline{\mathsf{B}}^3(r, x_0) \precsim \cup_{j=1}^{m} \overline{\mathsf{B}}^3(r, x_j) \precsim \overline{\mathsf{B}}^3(r, x_0) \cup \overline{\mathsf{B}}^3(r, x_0') \precsim \overline{\mathsf{B}}^3(r, x_0)$$

from the first part of the proof. From part (vii) of Lemma 5.5.44 we deduce that $\overline{\mathsf{B}}^3(r, x_0) \sim \cup_{j=1}^{m} \overline{\mathsf{B}}^3(r, x_j)$ as desired. ■

Now we may conclude the proof of the Banach–Tarski Paradox. Let X and Y be as in the statement of the theorem and let $x \in \mathrm{int}(X)$ and $y \in \mathrm{int}(Y)$. Choose $\epsilon \in \mathbb{R}_{>0}$ such that $\overline{\mathsf{B}}^3(\epsilon, x) \subseteq \mathrm{int}(X)$ and $\overline{\mathsf{B}}^3(\epsilon, y) \subseteq \mathrm{int}(Y)$. Boundedness of X ensures that there exists $x_1, \dots, x_k \in \mathbb{R}^3$ such that $X \subseteq \cup_{j=1}^{k} \overline{\mathsf{B}}^3(\epsilon, x_j)$. By Lemma 5.5.45 and part (vi) of Lemma 5.5.44 we have

$$\overline{\mathsf{B}}^3(\epsilon, \mathbf{0}) \precsim X \subseteq \cup_{j=1}^{k} \overline{\mathsf{B}}^3(\epsilon, x_j) \precsim \overline{\mathsf{B}}^3(\epsilon, \mathbf{0}).$$

By part (vii) of Lemma 5.5.44 it follows that $X \sim A$. Similarly we show that $Y \sim A$. By part (iii) of Lemma 5.5.44 it follows that $X \sim Y$, which is the result.

5.5.7 Notes

The Banach–Tarski Paradox is due to none other than **SB/AT:24**. The proof we give follows that of **KS:79**. The number 40 used in Lemma 5.5.43 is not optimal. Indeed, **TJD/JdG:56** show that one can decompose a ball into five disjoint components which can then be rearranged into two balls of the same size.

Exercises

5.5.1 Let $A \subseteq \mathbb{R}^n$ be Lebesgue measurable and for $\rho \in \mathbb{R}_{>0}$ define

$$\rho A = \{\rho \boldsymbol{x} \mid \boldsymbol{x} \in A\}.$$

Show that $\lambda_n(\rho A) = \rho^n \lambda_n(A)$.

5.5.2 Show that for $\boldsymbol{x} \in \mathbb{R}^n$, the point mass $\delta_{\boldsymbol{x}} \colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ is regular.

5.5.3 Show that the counting measure $\mu \colon \mathscr{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}_{\geq 0}$ is not regular.

Section 5.6

Measurable functions

In order to define the Lebesgue integral, one first defines functions for which it is *possible* to define the Lebesgue integral. What results is a quite general class of functions, certainly general enough to capture any function one is likely to encounter in that fantastic place called "The Real World."

Our approach, as with basic measure theory, is to start with generalities, and then proceed to particular aspects of Lebesgue measurable functions.

Do I need to read this section? If you are wanting to learn about integration in general, and the Lebesgue integral in particular, then this section is essential to this. •

5.6.1 General measurable maps and functions

We begin with a rather general definition of a measurable map between measurable spaces. The reader will observe that this definition harkens one back to the definition of continuity (cf. ***missing stuff***), and so can perhaps be seen as natural, provided you are comfortable with the naturality of continuity as in ***missing stuff***.

5.6.1 Definition (Measurable map) Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces. A map $f\colon X \to Y$ is ***($\mathscr{A}$, $\mathscr{B}$)-measurable*** if $f^{-1}(B) \in \mathscr{A}$ for every $B \in \mathscr{B}$. The set of $(\mathscr{A}, \mathscr{B})$-measurable maps is denoted by $\mathsf{L}^{(0)}((X, \mathscr{A}); (Y, \mathscr{B}))$, or simply by $\mathsf{L}^{(0)}(X; Y)$, with the understanding that the σ-algebras $\mathscr{A}$ and $\mathscr{B}$ are implicit. •

We shall not often consider maps between general measure spaces. However, the above general definition is useful because it gives some context for the particular definitions to follow.

It is often useful to be able to check measurability of a map by using generators for the σ-algebra involved. The following result is helpful for doing this.

5.6.2 Proposition (Measurability of maps using generators for σ-algebras) *Let* X *and* Y *be sets, let $\mathscr{S} \subseteq \mathbf{2}^{\mathrm{X}}$ and $\mathscr{T} \subseteq \mathbf{2}^{\mathrm{Y}}$, and let $\mathscr{A}_{\mathscr{S}}$ and $\mathscr{A}_{\mathscr{T}}$ be the σ-algebras generated by $\mathscr{S}$ and $\mathscr{T}$, respectively. If* $\mathrm{f}\colon \mathrm{X} \to \mathrm{Y}$ *is a map and if*

$$\mathscr{T} \subseteq \{\mathrm{T} \subseteq \mathrm{Y} \mid \mathrm{f}^{-1}(\mathrm{T}) \in \mathscr{S}\}$$

then f *is $(\mathscr{A}_{\mathscr{S}}, \mathscr{A}_{\mathscr{T}})$-measurable.*

Proof Let us denote

$$\mathscr{A}' = \{T \subseteq Y \mid f^{-1}(T) \in \mathscr{A}_{\mathscr{S}}\}.$$

We claim that $\mathscr{A}'$ is a σ-algebra containing $\mathscr{T}$. To see that it is a σ-algebra, first note that $f^{-1}(Y) = X \in \mathscr{A}_{\mathscr{S}}$ and so $Y \in \mathscr{A}'$. If $T \in \mathscr{A}'$ then

$$f^{-1}(Y \setminus T) = X \setminus f^{-1}(T)$$

by Exercise 1.3.3. Since $X \setminus f^{-1}(T) \in \mathscr{A}_{\mathscr{S}}$ by virtue of $\mathscr{A}_{\mathscr{S}}$ being a σ-algebra, it follows that $Y \setminus T \in \mathscr{A}'$. Finally, suppose that $(T_j)_{j\in\mathbb{Z}_{>0}}$ is a countable family of sets in $\mathscr{A}'$. Then, by Proposition 1.3.5 we have

$$f^{-1}\Big(\bigcup_{j\in\mathbb{Z}_{>0}} T_j\Big) = \bigcup_{j\in\mathbb{Z}_{>0}} f^{-1}(T_j) \in \mathscr{A}_{\mathscr{S}}$$

since $\mathscr{A}_{\mathscr{S}}$ is a σ-algebra. We thus conclude that $\cup_{j\in\mathbb{Z}_{>0}} T_j \in \mathscr{A}'$. This shows that $\mathscr{A}'$ is a σ-algebra. By hypothesis, if

$$B \in \mathscr{T} \subseteq \{T \subseteq Y \mid f^{-1}(T) \in \mathscr{S}\}$$

then $f^{-1}(B) \in \mathscr{S} \subseteq \mathscr{A}_{\mathscr{S}}$. Thus $\mathscr{T} \subseteq \mathscr{A}'$ and so $\mathscr{A}_{\mathscr{T}} \subseteq \mathscr{A}'$ since $\mathscr{A}_{\mathscr{T}}$ is the smallest σ-algebra containing $\mathscr{T}$. It therefore follows that if $B \in \mathscr{A}_{\mathscr{T}}$ then $f^{-1}(B) \in \mathscr{A}_{\mathscr{S}}$, i.e., that f is $(\mathscr{A}_{\mathscr{S}}, \mathscr{A}_{\mathscr{T}})$-measurable. ■

We can give an application of the preceding result that gives an important class of measurable maps.

5.6.3 Example (Continuous maps are Borel-measurable) We claim that if $\boldsymbol{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous then it is $(\mathscr{B}(\mathbb{R}^n), \mathscr{B}(\mathbb{R}^m))$-measurable. Indeed, $\mathscr{B}(\mathbb{R}^n)$ and $\mathscr{B}(\mathbb{R}^m)$ are the σ-algebras generated by the collections $\mathscr{O}(\mathbb{R}^n)$ and $\mathscr{O}(\mathbb{R}^m)$ of open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. Since $\boldsymbol{f}$ is continuous it follows from Corollary **??** that

$$\mathscr{O}(\mathbb{R}^m) \subseteq \{U \subseteq \mathbb{R}^m \mid \boldsymbol{f}^{-1}(U) \in \mathscr{O}(\mathbb{R}^n)\}.$$

From Proposition 5.6.2 we conclude that $\boldsymbol{f}$ is $(\mathscr{B}(\mathbb{R}^n), \mathscr{B}(\mathbb{R}^m))$-measurable. •

What we are really interested in in this section are $\mathbb{R}$-valued functions. It turns out to be interesting to consider $\overline{\mathbb{R}}$-valued functions. The reason for this degree of generality is not that we are interested in infinite-valued functions *per se*, but that we are interested in sequences of $\mathbb{R}$-valued functions that turn out to have infinite limits. The reader will want to be familiar with the order relations on $\overline{\mathbb{R}}$ defined in Section 2.2.5.

In any case, we now turn our attention to functions $f\colon X \to \overline{\mathbb{R}}$ defined on a measurable space $(X, \mathscr{A})$. For such functions we have the following equivalent properties.

5.6.4 Proposition (Characterisations of measurable functions) *For a measurable space* $(\mathrm{X}, \mathscr{A})$ *and a map* $\mathrm{f}\colon \mathrm{X} \to [-\infty, \infty]$*, the following statements are equivalent:*

(i) *for each* $\mathrm{b} \in \mathbb{R}$ *the set* $\mathrm{f}^{-1}([-\infty, \mathrm{b}]) = \{\mathrm{x} \in \mathrm{X} \mid \mathrm{f}(\mathrm{x}) \leq \mathrm{b}\}$ *is measurable;*

(ii) *for each* $\mathrm{b} \in \mathbb{R}$ *the set* $\mathrm{f}^{-1}([-\infty, \mathrm{b})) = \{\mathrm{x} \in \mathrm{X} \mid \mathrm{f}(\mathrm{x}) < \mathrm{b}\}$ *is measurable;*

(iii) *for each* $\mathrm{a} \in \mathbb{R}$ *the set* $\mathrm{f}^{-1}([\mathrm{a}, \infty]) = \{\mathrm{x} \in \mathrm{X} \mid \mathrm{f}(\mathrm{x}) \geq \mathrm{a}\}$ *is measurable;*

(iv) *for each* $\mathrm{a} \in \mathbb{R}$ *the set* $\mathrm{f}^{-1}((\mathrm{a}, \infty]) = \{\mathrm{x} \in \mathrm{X} \mid \mathrm{f}(\mathrm{x}) > \mathrm{a}\}$ *is measurable.*

Proof (i) $\Longrightarrow$ (ii) We write

$$f^{-1}([-\infty, b)) = f^{-1}(\cup_{k\in\mathbb{Z}_{>0}} f^{-1}([-\infty, b - \tfrac{1}{k}]) = \cup_{k\in\mathbb{Z}_{>0}} f^{-1}([-\infty, b - \tfrac{1}{k}])$$

by Proposition 1.3.5. Since $f^{-1}([-\infty, b-\frac{1}{k}] \in \mathscr{A}$ by assumption and since $\mathscr{A}$ is a σ-algebra, we conclude that $f^{-1}([-\infty, b)) \in \mathscr{A}$.

(ii) $\Longrightarrow$ (iii) Here we note that

$$f^{-1}([a, \infty]) = X \setminus f^{-1}([-\infty, a))$$

by Exercise 1.3.3. Since $\mathscr{A}$ is a σ-algebra and since $f^{-1}([-\infty, a)) \in \mathscr{A}$ by assumption, it follows that $f^{-1}([a, \infty]) \in \mathscr{A}$.

(iii) $\Longrightarrow$ (iv) Here we write

$$f^{-1}((a, \infty]) = \cup_{k\in\mathbb{Z}_{>0}} f^{-1}([a+\tfrac{1}{k}, \infty]) =$$

by Proposition 1.3.5. As in the first part of the proof we conclude that $f^{-1}((a, \infty]) \in \mathscr{A}$.

(iv) $\Longrightarrow$ (i) Here we note that

$$f^{-1}([-\infty, b]) = X \setminus f^{-1}((b, \infty])$$

by Exercise 1.3.3 and then argue as in the second part of the proof that $f^{-1}([-\infty, b]) \in \mathscr{A}$. ■

With this result at hand, the following definition makes sense.

5.6.5 Definition (Measurable function) For a measurable space $(X, \mathscr{A})$ a function $f\colon X \to \overline{\mathbb{R}}$ satisfying any one of the four equivalent conditions of Proposition 5.6.4 is an ***$\mathscr{A}$-measurable*** function. We shall frequently just say that f is ***measurable*** if $\mathscr{A}$ is understood. For any subset $I \subseteq \overline{\mathbb{R}}$ (typically we will be concerned with $I \in \{\mathbb{R}, \overline{\mathbb{R}}_{\geq 0}\}$) we denote the set of measurable I-valued maps by $\mathsf{L}^{(0)}((X, \mathscr{A}); I)$, or by $\mathsf{L}^{(0)}(X; I)$, with the understanding that the σ-algebra $\mathscr{A}$ is implicit. •

The relationship of this notion of measurability with that of Definition 5.6.1 is perhaps not immediately clear. So let us make this clear, recalling from Definition 5.4.15 the definition of the σ-algebra $\mathscr{B}(\overline{\mathbb{R}})$ on $\overline{\mathbb{R}}$.

5.6.6 Proposition (Characterisation of measurable functions) *For a measurable space $(X, \mathscr{A})$ and a map $f\colon X \to \overline{\mathbb{R}}$, the following statements are equivalent:*

(i) $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$;

(ii) the sets $\{x \in X \mid f(x) = -\infty\}$ and $\{x \in X \mid f(x) = \infty\}$ are measurable and $f^{-1}(B) \in \mathscr{A}$ for every $B \in \mathscr{B}(\mathbb{R})$;

(iii) f is $(\mathscr{A}, \mathscr{B}(\overline{\mathbb{R}}))$-measurable.

Proof (i) $\Longrightarrow$ (ii) We have

$$f^{-1}(-\infty) = f^{-1}(\cap_{k\in\mathbb{Z}_{>0}}[-\infty, -k]) = \cap_{k\in\mathbb{Z}} f^{-1}([-\infty, -k]),$$
$$f^{-1}(\infty) = f^{-1}(\cap_{k\in\mathbb{Z}_{>0}}[k, \infty]) = \cap_{k\in\mathbb{Z}} f^{-1}([k, \infty]),$$

by Proposition 1.3.5. Thus $f^{-1}(-\infty)$ and $f^{-1}(\infty)$ are countable intersections of measurable sets and so themselves measurable. We must also show that $f^{-1}(B)$ is measurable for a Borel set B. To prove this, we denote

$$\mathscr{B}'(\mathbb{R}) = \{S \subseteq \mathbb{R} \mid f^{-1}(S) \in \mathscr{A}\}.$$

We claim that $\mathscr{B}'(\mathbb{R})$ is a σ-algebra containing $\mathscr{B}(\mathbb{R})$. Certainly $\mathbb{R} \in \mathscr{B}'(\mathbb{R})$ since $f^{-1}(\mathbb{R}) = X \in \mathscr{A}$. If $(S_j)_{j\in\mathbb{Z}_{>0}}$ is a countable collection of subsets from $\mathscr{B}'(\mathbb{R})$ we have

$$f^{-1}\Big(\bigcup_{j\in\mathbb{Z}_{>0}} S_j\Big) = \bigcup_{j\in\mathbb{Z}_{>0}} f^{-1}(S_j) \in \mathscr{A},$$

where we have used Proposition 1.3.5. Thus $\cup_{j\in\mathbb{Z}_{>0}} S_j \in \mathscr{B}'(\mathbb{R})$. Also, by Exercise 1.3.3, if $S \in \mathscr{B}'(\mathbb{R})$ then

$$f^{-1}(\mathbb{R} \setminus S) = X \setminus f^{-1}(S) \in \mathscr{A}$$

and so $\mathbb{R}\setminus S \in \mathscr{B}'(\mathbb{R})$. Thus $\mathscr{B}'(\mathbb{R})$ is a σ-algebra. By hypothesis we have $(-\infty, b] \in \mathscr{B}'(\mathbb{R})$ for every $b \in \mathbb{R}$. Thus $\mathscr{B}'(\mathbb{R})$ contains the σ-algebra generated by sets of the form $(-\infty, b]$ for $b \in \mathbb{R}$. By Proposition 5.4.9 this means that $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{B}'(\mathbb{R})$, as claimed. This proves that $f^{-1}(B) \in \mathscr{A}$ for $B \in \mathscr{B}(\mathbb{R})$.

(ii) $\Longrightarrow$ (iii) Let

$$\mathscr{B}'(\overline{\mathbb{R}}) = \{T \subseteq \overline{\mathbb{R}} \mid f^{-1}(T) \in \mathscr{A}\},$$

and note that, by hypothesis, $\mathscr{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\} \subseteq \mathscr{B}'(\overline{\mathbb{R}})$. By Proposition 5.4.16 it follows that f is $(\mathscr{A}, \mathscr{B}(\overline{\mathbb{R}}))$-measurable.

(iii) $\Longrightarrow$ (i) For $a \in \mathbb{R}$ we have

$$f^{-1}((a,\infty]) = f^{-1}((a,\infty) \cup \{\infty\}) = f^{-1}((a,\infty)) \cup f^{-1}(\{\infty\})$$

by Proposition 1.3.5. Since (a,∞) is open it is a Borel set and so in $\mathscr{B}(\overline{\mathbb{R}})$ by Proposition 5.4.16. Thus $f^{-1}((a,\infty)) \in \mathscr{A}$ by hypothesis. Also, $\{\infty\} \in \mathscr{B}(\overline{\mathbb{R}})$ by Proposition 5.4.16 and so $f^{-1}(\{\infty\}) \in \mathscr{A}$. Therefore, $f^{-1}((a,\infty])$ is a union of measurable sets and so is measurable. Thus f is $\mathscr{A}$-measurable. ■

For functions that are $\mathbb{R}$-valued this gives the following result.

5.6.7 Corollary (Measurability of $\mathbb{R}$-valued functions) *For a measurable space* $(X, \mathscr{A})$, *a function* $f\colon X \to \mathbb{R}$ *is measurable if and only if it is* $(\mathscr{A}, \mathscr{B}(\mathbb{R}))$*-measurable.*

It is often fairly easy to apply Definition 5.6.5 to ascertain whether a given function is measurable (as opposed to employing the equivalent characterisation of Proposition 5.6.6).

5.6.8 Examples (Measurable functions)

1. For a measurable space $(X, \mathscr{A})$ and for $\alpha \in \overline{\mathbb{R}}$, we claim that the constant function $f_\alpha\colon x \mapsto \alpha$ is $\mathscr{A}$-measurable. To see this we let $b \in \mathbb{R}$ and determine that

$$f_\alpha^{-1}([-\infty, b)) = \begin{cases} \emptyset, & b \le \alpha, \\ X, & b > \alpha, \end{cases}$$

provided that $\alpha \neq -\infty$. If $\alpha = -\infty$ then $f_\alpha^{-1}([-\infty, b)) = X$ for every $b \in \mathbb{R}$. In any case, $f_\alpha^{-1}([-\infty, b)) \in \mathscr{A}$ for all $b \in \mathbb{R}$ and so f_α is $\mathscr{A}$-measurable.

2. Let $(X, \mathscr{A})$ be a measurable space and let $A \in \mathscr{A}$. We claim that the characteristic function $\chi_A \colon X \to \mathbb{R}$ is $\mathscr{A}$-measurable. Indeed,

$$\chi_A^{-1}([a, \infty]) = \begin{cases} X, & a \leq 0, \\ A, & a \in (0, 1], \\ \emptyset, & a > 1. \end{cases}$$

Since $X, A, \emptyset \in \mathscr{A}$ it follows that χ_A is indeed $\mathscr{A}$-measurable.

Note that the same argument shows that, if $A \notin \mathscr{A}$, then χ_A is not $\mathscr{A}$-measurable.

3. Let $A \in \mathscr{L}(\mathbb{R}^n)$ and let $f \colon A \to \mathbb{R}$ be continuous. We claim that f is $\mathscr{L}(\mathbb{R}^n)$-measurable. Indeed, for $a \in \mathbb{R}$ the set $f^{-1}((a, \infty))$ is open in A by Corollary **??**. Thus there exists an open subset $U_a \subseteq \mathbb{R}^n$ such that $f^{-1}((a, \infty)) = U_a \cap A$. Since $U_a \in \mathscr{L}(\mathbb{R}^n)$ (open sets are Borel sets and so are Lebesgue measurable) we have $f^{-1}((a, \infty)) \in \mathscr{L}(\mathbb{R}^n)$ and so is a measurable subset of A. •

Let $(X, \mathscr{A})$ be a measurable space. By Corollary 5.6.7, measurability of $f \colon X \to \mathbb{R}$ is equivalent to $(\mathscr{A}, \mathscr{B}(\mathbb{R}))$-measurability of f. A natural question to ask is: "Why use the σ-algebra of *Borel* sets on $\mathbb{R}$ to define measurability of a function? Why not use the σ-algebra of *Lebesgue* measurable sets?" The answer to this question perhaps cannot be divined immediately. The reason for using the Borel measurable sets is answered by answering the question, "What is it we are trying to achieve with our definition of a measurable function?" We shall not address this here, but instead refer ahead to Section 5.6.5.***missing stuff*** For now, let us simply illustrate that $(\mathscr{A}, \mathscr{B}(\mathbb{R}))$-measurability and $(\mathscr{A}, \mathscr{L}(\mathbb{R}))$-measurability are not equivalent.

5.6.9 Example ($(\mathscr{A}, \mathscr{B}(\mathbb{R}))$- and $(\mathscr{A}, \mathscr{L}(\mathbb{R}))$-measurability are different) Since $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{L}(\mathbb{R})$ it follows that $f \colon X \to \mathbb{R}$ is $(\mathscr{A}, \mathscr{B}(\mathbb{R}))$-measurable if it is $(\mathscr{A}, \mathscr{L}(\mathbb{R}))$-measurable. The converse implication is not generally true, however. We illustrate this with an example. We take $X = [0, 1]$ and $\mathscr{A} = \mathscr{L}([0, 1])$. We define a function $f \colon [0, 1] \to \mathbb{R}$ that is $(\mathscr{L}([0, 1]), \mathscr{B}(\mathbb{R}))$-measurable but not $(\mathscr{L}([0, 1]), \mathscr{L}(\mathbb{R}))$-measurable. Our construction relies on the reader understanding the construction of the sets C_ϵ and C from Examples 2.5.42 and 2.5.39.

Let $\epsilon \in \mathbb{R}_{>0}$ and let $C_\epsilon \subseteq [0, 1]$ be the "fat" Cantor set of Example 2.5.42. Let $C \subseteq [0, 1]$ be the standard middle-thirds Cantor set of Example 2.5.39. Recall that the inductive construction of these sets is the same in that they are defined by, at step k, removing 2^k open intervals from the set defined at step $k - 1$. This defines countable collections $(I_{\epsilon,k})_{k \in \mathbb{Z}_{>0}}$ and $(I_k)_{k \in \mathbb{Z}_{>0}}$ of disjoint open intervals such that

$$C_\epsilon = [0, 1] \setminus \cup_{j \in \mathbb{Z}_{>0}} I_{\epsilon,j}, \quad C = [0, 1] \setminus \cup_{j \in \mathbb{Z}_{>0}} I_j.$$

Moreover, since the constructions of C_ϵ and C proceed in the same way, the intervals $(I_{\epsilon,j})_{j \in \mathbb{Z}_{>0}}$ and $(I_j)_{j \in \mathbb{Z}_{>0}}$ can be enumerated consistently such $I_{\epsilon,1}$ and I_1 are the intervals removed in the first step in the inductive constructions of C_ϵ and C, $I_{\epsilon,2}$ and $I_{\epsilon,3}$, and I_2 and I_3 are the intervals, ordered from left to right, removed in the second step in the inductive constructions of C_ϵ and C, and so on. We then define $f \colon [0, 1] \to \mathbb{R}$ by asking that $f|I_{\epsilon,j}$ maps $I_{\epsilon,j}$ linearly onto the interval I_j, mapping the left (resp. right) endpoint of $I_{\epsilon,j}$ to the left (resp. right) endpoint of I_j. Note that since $\mathrm{cl}([0,1] \setminus C_\epsilon) = [0,1]$, it follows that this definition of f on $[0,1] \setminus C_\epsilon$ extends to a continuous function f from $[0,1]$ to $\mathbb{R}$. By Example 5.6.8–3 it follows that f is $(\mathscr{L}([0,1]), \mathscr{B}(\mathbb{R}))$-measurable. Moreover, $f(C_\epsilon) = C$ since, by construction, the points in C_ϵ and C are the endpoints of intervals from $(I_{\epsilon,j})_{j\in\mathbb{Z}_{>0}}$ and $(I_j)_{j\in\mathbb{Z}_{>0}}$. Moreover, we claim that f is strictly monotonically increasing. It is obviously monotonically increasing. To see that it is strictly monotonically increasing, suppose that $x_1, x_2 \in [0,1]$ satisfy $x_1 < x_2$ and $f(x_1) = f(x_2)$. This means that $f|[x_1, x_2]$ is constant which, by construction of f implies that $[x_1, x_2] \subseteq C_\epsilon$, contradicting the fact that $\mathrm{int}(C_\epsilon) = \emptyset$. Thus f is strictly monotonically increasing and so injective by Theorem 3.1.30. By Theorem 5.4.5 there exists a subset $S \subseteq C_\epsilon$ that is not Lebesgue measurable. Let $T = f(S) \subseteq C$ so that $T \in \mathscr{L}(\mathbb{R})$ since $\lambda(C) = 0$ and since $\mathscr{L}(\mathbb{R})$ is complete. Injectivity of f implies that $f^{-1}(T) = S \notin \mathscr{L}([0,1])$. Thus f is not $(\mathscr{L}([0,1]), \mathscr{L}(\mathbb{R}))$-measurable. •

It is often the case that one is able to draw conclusions about a function only almost everywhere, not everywhere. In such cases, one would like to assert that this almost everywhere knowledge of the function is enough to ensure its measurability. It should not be surprising that completeness plays a rôle here. Note that this is the first time we have used a measure in our discussion of measurable functions. Up to now we have only used measurable spaces.

5.6.10 Proposition (Measurability of almost everywhere known functions) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a complete measure space, if* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *is* $\mathscr{A}$*-measurable, and if* $\mathsf{g}\colon \mathsf{X} \to \overline{\mathbb{R}}$ *satisfies*

$$\mu(\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) \neq \mathsf{g}(\mathsf{x})\}) = 0,$$

then $\mathsf{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$.

Proof Let

$$A_{f,g} = \{x \in A \mid f(x) = g(x)\}$$

and let $b \in \mathbb{R}$. Then

$$\{x \in X \mid g(x) \leq b\} = (\{x \in X \mid f(x) \leq b\} \cap A_{f,g}) \cup (\{x \in X \mid g(x) \leq b\} \cap (X \setminus A_{f,g})).$$

The set $X \setminus A_{f,g}$ has measure zero and so is measurable. Thus $A_{f,g}$ is measurable and so the set

$$\{x \in X \mid f(x) \leq b\} \cap A_{f,g}$$

is measurable. Since the set

$$\{x \in X \mid g(x) \leq b\} \cap (X \setminus A_{f,g})$$

is a subset of the set $A_{f,g}$ which has measure zero, completeness of $(X, \mathscr{A}, \mu)$ ensures that it has measure zero, and in particular is measurable. Thus

$$\{x \in X \mid g(x) \leq b\}$$

is the intersection of measurable sets, and so is measurable. ■

5.6.2 Measurability and operations on functions

At this point we are still not clear on the significance of measurable functions, and we will continue to postpone this until Section 5.6.5. All we really know at the moment is that the set of measurable functions on $(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n))$ contains the continuous functions, and so there is a nice subset of measurable functions in this case. It turns out that measurable functions also have nice properties with respect to the natural operations one performs on functions and sequences of functions. In this section we prove these properties.

We begin with the interaction of measurable functions with standard algebraic operations. In order to do this, the reader will wish to recall from Section 2.2.5 the "algebraic" operations on $\overline{\mathbb{R}}$. This is complicated a little for measurable functions since these are $\overline{\mathbb{R}}$-valued. To properly state the result we need, it is, therefore, convenient to introduce some notation to account for the fact that certain algebraic operations are ill-defined on $\overline{\mathbb{R}}$. If X is a set, if $f\colon X \to \overline{\mathbb{R}}$, and if $\alpha_-, \alpha_+ \in \overline{\mathbb{R}}$, then we denote by $f_{\alpha_-,\alpha_+}\colon X \to \overline{\mathbb{R}}$ the function given by

$$f_{\alpha_-,\alpha_+}(x) = \begin{cases} f(x), & f(x) \in \mathbb{R}, \\ \alpha_-, & f(x) = -\infty, \\ \alpha_+, & f(x) = \infty. \end{cases}$$

Similarly, for $\alpha_-, \alpha_+, \alpha_0 \in \mathbb{R}$ we denote by $f_{\alpha_-,\alpha_+,\alpha_0}\colon X \to \mathbb{R}$ the function given by

$$f_{\alpha_-,\alpha_+,\alpha_0}(x) = \begin{cases} \alpha_-, & f(x) = -\infty, \\ \alpha_+, & f(x) = \infty, \\ \alpha_0, & f(x) = 0, \\ f(x), & \text{otherwise.} \end{cases}$$

Next, for $f, g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$ and for $\alpha \in \overline{\mathbb{R}}$ denote $f +_\alpha g\colon X \to \overline{\mathbb{R}}$ the function defined by

$$(f +_\alpha g)(x) = \begin{cases} \alpha, & f(x) = \infty,\ g(x) = -\infty \text{ or } f(x) = -\infty,\ g(x) = \infty, \\ f(x) + g(x), & \text{otherwise.} \end{cases}$$

With these tedious bits of notation out of the way, we can now state the desired result.

5.6.11 Proposition (Algebraic operations on measurable functions) *Let* $(X, \mathscr{A})$ *be a measurable space, let* $f, g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$*, let* $\beta \in \mathbb{R}$*, let* $\beta_-, \beta_+, \beta_0 \in \mathbb{R}^*$*, let* $\alpha, \alpha_-, \alpha_+ \in \overline{\mathbb{R}}$*, let* $p \in \mathbb{R}_{>0}$*, and let* $k \in \mathbb{Z}_{>0}$*. Then the following functions are* $\mathscr{A}$*-measurable:*

(i) βf;

(ii) $f +_\alpha g$;

(iii) fg;

(iv) $\dfrac{f}{g_{\beta_-,\beta_+,\beta_0}}$;

(v) $(|f|^p)_{\alpha_-,\alpha_+}$;

(vi) $(f^k)_{\alpha_-,\alpha_+}$.

Proof We shall freely make use of Proposition 5.6.13 below.

(i) Let $\phi_\beta \colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ be defined by $\phi_\beta(y) = \beta y$. Then $\beta f = \phi_\beta \circ f$. Since

$$\phi_\beta^{-1}(U) = \{\beta y \mid y \in U\}$$

it follows that $\phi_\beta^{-1}(U)$ is open for open set U. Also,

$$\phi_\beta^{-1}([-\infty, b)) = \begin{cases} [-\infty, \beta b), & \beta \in \mathbb{R}_{>0}, \\ \{0\}, & \beta = 0, \\ (\beta b, \infty], & \beta \in \mathbb{R}_{<0}, \end{cases}$$

and so $\phi_\beta^{-1}([-\infty, b)) \in \mathscr{B}(\overline{\mathbb{R}})$ for every $\beta, b \in \mathbb{R}$. Similarly, $\phi_\beta^{-1}((a, \infty]) \in \mathscr{B}(\overline{\mathbb{R}})$ for every $\beta, a \in \mathbb{R}$. From this we deduce, using Proposition 1.3.5, that the preimage by ϕ_β of the generators of the σ-algebra $\mathscr{B}(\overline{\mathbb{R}})$ are in $\mathscr{B}(\overline{\mathbb{R}})$. By Proposition 5.6.2 we conclude that ϕ_β is $\mathscr{B}(\overline{\mathbb{R}})$-measurable. By Proposition 5.6.13 we then conclude that βf is $\mathscr{A}$-measurable.

(ii) Here we use a pair of fairly simple lemmata.

1 **Lemma** *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and* $\mathscr{A}$*-measurable functions* $\mathsf{f}, \mathsf{g} \colon \mathsf{X} \to \overline{\mathbb{R}}$*, the following sets are measurable:*

(i) $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) > \mathsf{g}(\mathsf{x})\}$;

(ii) $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) \geq \mathsf{g}(\mathsf{x})\}$;

(iii) $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) = \mathsf{g}(\mathsf{x})\}$.

Proof (i) We claim that

$$\{x \in X \mid f(x) > g(x)\} = \bigcup_{q \in \mathbb{Q}} \big(\{x \in X \mid f(x) > q\} \cap \{x \in X \mid g(x) < q\}\big).$$

Indeed, let $x \in \{x' \in X \mid f(x') > g(x')\}$. If $f(x) = \infty$ then $g(x) < \infty$. Thus there exists $q \in \mathbb{Q}$ such that $f(x) > q$ and $g(x) < q$. If $f(x) < \infty$ then $f(x) \in \mathbb{R}$ since we cannot have $f(x) = -\infty$. Therefore, there exists $q \in \mathbb{Q}$ such that $f(x) > q$ and $g(x) < q$. This shows that

$$\{x \in X \mid f(x) > g(x)\} \subseteq \bigcup_{q \in \mathbb{Q}} \big(\{x \in X \mid f(x) > q\} \cap \{x \in X \mid g(x) < q\}\big).$$

For the converse inclusion, suppose that $x \in X$ has the property that there exists $q \in \mathbb{Q}$ such that $g(x) < q < f(x)$. Clearly $x \in \{x' \in X \mid f(x') > g(x')\}$, giving our claim.

Now, since f and g are $\mathscr{A}$-measurable, the sets

$$\{x \in X \mid f(x) > q\}, \quad \{x \in X \mid g(x) < q\}, \qquad q \in \mathbb{Q},$$

are measurable, and so too then is their intersection. Thus $\{x \in X \mid f(x) > g(x)\}$ is a countable union of measurable sets, which is then measurable.

(ii) Note that

$$\{x \in X \mid f(x) \geq g(x)\} = X \setminus \{x \in X \mid g(x) > f(x)\}.$$

Since $\{x \in X \mid g(x) > f(x)\}$ is measurable by the first part of the lemma it follows that $\{x \in X \mid f(x) \geq g(x)\}$ is also measurable.

(iii) We have

$$\{x \in X \mid f(x) = g(x)\} = \{x \in X \mid f(x) \geq g(x)\} \cap \{x \in X \mid g(x) \geq f(x)\}.$$

The right-hand side is the intersection of two measurable sets by the second part of the lemma, and so is measurable. ▼

2 Lemma *If* $(\mathsf{X}, \mathscr{A})$ *is a measurable space, if* $\mathsf{f}\colon \mathsf{X} \to \overline{\mathbb{R}}$ *is* $\mathscr{A}$*-measurable, and if* $\beta \in \mathbb{R}$*, then the function* $\mathsf{x} \mapsto \mathsf{f}(\mathsf{x}) + \beta$ *is* $\mathscr{A}$*-measurable.*

Proof Define $\phi_\beta\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ by $\phi_\beta(y) = y + \beta$. By Proposition 5.4.22 it follows that $\phi_\beta^{-1}(B) \in \mathscr{B}(\mathbb{R})$ for $B \in \mathscr{B}(\mathbb{R})$. It is clear that $\phi_\beta^{-1}(\{-\infty\}) = \{-\infty\}$ and that $\phi_\beta^{-1}(\{\infty\}) = \{\infty\}$. Therefore, by Propositions 5.4.16 and 5.6.2, it follows that ϕ_β is $\mathscr{B}(\overline{\mathbb{R}})$-measurable. Thus $\phi_\beta \circ f$ is $\mathscr{A}$-measurable by Proposition 5.6.13. ▼

To proceed with the proof, let $a \in \mathbb{R}$ and let

$$A_{a,\alpha} = (\{x \in X \mid f(x) = \infty\} \cap \{x \in X \mid g(x) = -\infty\}) \\ \cup (\{x \in X \mid f(x) = -\infty\} \cap \{x \in X \mid g(x) = \infty\})$$

if $a < \alpha$ and let $A_{a,\alpha} = \emptyset$ if $a \geq \alpha$. We then have

$$\begin{aligned}(f +_\alpha g)^{-1}((a, \infty]) &= \{x \in X \mid f(x) + g(x) > a\} \cup A_{a,\alpha} \\ &= \{x \in X \mid f(x) > a - g(x)\} \cup A_{a,\alpha}.\end{aligned}$$

By the two lemmata above, the set $\{x \in X \mid f(x) > a - g(x)\}$ is measurable. By Proposition 5.6.6 each of the four sets comprising the definition of $A_{a,\alpha}$ when $a < \alpha$ is measurable. Thus $A_{a,\alpha}$ is measurable and so $(f +_\alpha g)^{-1}((a, \infty])$ is measurable, being a union of measurable sets.

(iii) We denote

$$\begin{aligned}A_{f,-} &= \{x \in X \mid f(x) = -\infty\}, \quad & A_{f,+} &= \{x \in X \mid f(x) = \infty\}, \\ A_{g,-} &= \{x \in X \mid g(x) = -\infty\}, \quad & A_{g,+} &= \{x \in X \mid g(x) = \infty\}.\end{aligned}$$

By Proposition 5.6.4 these sets are measurable. For $x \notin A_{f,-} \cup A_{f,+} \cup A_{g,-} \cup A_{g,+}$ we have

$$f(x)g(x) = \tfrac{1}{2}((f(x) + g(x))^2 - f(x)^2 - g(x)^2).$$

If $x \in A_{f,-} \cap A_{g,+}$ or $x \in A_{f,+} \cap A_{g,-}$ then $f(x)g(x) = -\infty$ and if $x \in A_{f,-} \cap A_{g,-}$ or $x \in A_{f,+} \cap A_{g,+}$ then $f(x)g(x) = \infty$. Then, for $a \in \mathbb{R}$ we have

$$\begin{aligned}(fg)^{-1}((a, \infty]) &= \{x \in X \mid f(x)g(x) > a\} \\ &= \{x \in (A_{f,-} \cap A_{g,+}) \cup (A_{f,+} \cap A_{g,-}) \mid f(x)g(x) > a\} \\ &\quad \cup \{x \in (A_{f,-} \cap A_{g,-}) \cup (A_{f,+} \cap A_{g,+}) \mid f(x)g(x) > a\} \\ &\quad \cup \{x \in X \setminus (A_{f,-} \cup A_{f,+} \cup A_{g,-} \cup A_{g,+}) \mid \\ &\quad \tfrac{1}{2}((f(x) + g(x))^2 - f(x)^2 - g(x)^2) > a\}.\end{aligned}$$

The set

$$\{x \in (A_{f,-} \cap A_{g,+}) \cup (A_{f,+} \cap A_{g,-}) \mid f(x)g(x) > a\}$$

is empty, the set

$$\{x \in (A_{f,-} \cap A_{g,-}) \cup (A_{f,+} \cap A_{g,+}) \mid f(x)g(x) > a\}$$

is measurable being a union of measurable sets, and the set

$$\{x \in X \setminus (A_{f,-} \cup A_{f,+} \cup A_{g,-} \cup A_{g,+}) \mid \tfrac{1}{2}((f(x)+g(x))^2 - f(x)^2 - g(x)^2) > a\}$$

is measurable by parts (ii) and (vi). Thus $(fg)^{-1}((a,\infty])$ is a union of three measurable sets and so measurable.

(iv) We first consider the case when $f(x) = 1$ for every $x \in X$. In this case let us define $\phi_{\beta_-,\beta+,\beta_0}\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ by

$$\phi_{\beta_-,\beta+,\beta_0}(y) = \begin{cases} \frac{1}{y}, & y \in \mathbb{R}, \\ \frac{1}{\beta_0}, & y = 0, \\ \frac{1}{\beta_-}, & y = -\infty, \\ \frac{1}{\beta_+}, & y = \infty. \end{cases}$$

Note that $y \mapsto \frac{1}{y}$ is $(\mathscr{B}(\mathbb{R}), \mathscr{B}(\mathbb{R}))$-measurable by Example 5.6.3. Therefore, by Propositions 5.4.16 and 5.6.2 it is easy to see that $\phi_{\beta_-,\beta+,\beta_0}$ is $\mathscr{B}(\overline{\mathbb{R}})$-measurable. Since $\frac{1}{g_{\beta_-,\beta+,\beta_0}} = \phi_{\beta_-,\beta+,\beta_0} \circ g$ this part of the result follows from Proposition 5.6.13 in the case that $f = 1$. For general f the result follows from the result for $f = 1$ and from part (iii).

(v) Here we define $\phi_{\alpha_-,\alpha_+}\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ by

$$\phi_{\alpha_-,\alpha_+}(y) = \begin{cases} |y|^p, & y \in \mathbb{R}, \\ \alpha_-, & y = -\infty, \\ \alpha_+, & y = \infty. \end{cases}$$

It is easy to verify by Propositions 5.4.16 and 5.6.2 that ϕ_{α_-,α_+} is $\mathscr{B}(\overline{\mathbb{R}})$-measurable. Since the function $y \mapsto |y|^p$ is continuous and so $(\mathscr{B}(\mathbb{R}), \mathscr{B}(\mathbb{R}))$-measurable by Example 5.6.3. Thus, since $(|f|^p)_{\alpha_-,\alpha_+} = \phi_{\alpha_-,\alpha_+} \circ f$, this part of the result follows from Proposition 5.6.13.

(vi) If we define $\phi_{\alpha_-,\alpha_+}\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ as in the proof of part (v), this part of the proof is carried out exactly as that for part (v). ■

We now consider the interaction of composition and measurability. First of all, the most general result is false as the following example shows.

5.6.12 Example (Compositions of measurable functions may not be measurable) We recall from Exercise 5.6.9 the construction of a map $f\colon [0,1] \to [0,1]$ that is $(\mathscr{L}([0,1]), \mathscr{B}(\mathbb{R}))$-measurable but not $(\mathscr{L}([0,1]), \mathscr{L}(\mathbb{R}))$-measurable. Let $S, T \subseteq [0,1]$ be the subsets constructed in Exercise 5.6.9 and let $\chi_T\colon [0,1] \to \mathbb{R}$ be the characteristic function. Since T is Lebesgue measurable, as we showed in Exercise 5.6.9, it follows from Example 5.6.8–2 that χ_T is $\mathscr{L}([0,1])$-measurable. However, by construction of f, $\chi_T \circ f = \chi_S$. Since $S \notin \mathscr{L}([0,1])$ by construction, it follows from Example 5.6.8–2 that χ_S is not $\mathscr{L}([0,1])$-measurable. Thus the composition of measurable functions need not be measurable. •

The preceding counterexample notwithstanding, there is a useful result concerning measurability of compositions. The result relies on the notion of the σ-algebra $\mathscr{B}(\overline{\mathbb{R}})$ on $\overline{\mathbb{R}}$ as defined in Definition 5.4.15.

5.6.13 Proposition (Composition and measurable functions) *Let* $(\mathsf{X},\mathscr{A})$ *be a measurable space, let* $\mathsf{f}\in\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}})$*, and let* $\phi\colon\overline{\mathbb{R}}\to\overline{\mathbb{R}}$ *be* $\mathscr{B}(\overline{\mathbb{R}})$*-measurable. Then* $\phi\circ\mathsf{f}\in\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}})$*.*

Proof Let $B\in\mathscr{B}(\overline{\mathbb{R}})$. By assumption and by Proposition 5.6.6 we have $\phi^{-1}(B)\in\mathscr{B}(\overline{\mathbb{R}})$. Thus, using Exercise 1.3.2,

$$(\phi\circ f)^{-1}(B)=f^{-1}(\phi^{-1}(B))\in\mathscr{A},$$

and so $\phi\circ f$ is $\mathscr{A}$-measurable, as desired. ■

5.6.14 Corollary (Composition by continuous functions and measurability) *Let* $(\mathsf{X},\mathscr{A})$ *be a measurable space, let* $\mathsf{f}\in\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}})$*, and let* $\phi\colon\overline{\mathbb{R}}\to\overline{\mathbb{R}}$ *be continuous. Then* $\phi\circ\mathsf{f}\in\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}})$*.*

Proof This follows from Proposition 5.6.13, along with Example 5.6.3. ■

In the following result we consider measurability of functions restricted to measurable sets. We recall from Proposition 5.2.6 the definition of the restriction $\mathscr{A}_A$ of a measurable space $(X,\mathscr{A})$ to a measurable subset $A\in\mathscr{A}$.

5.6.15 Proposition (Measurability and restriction) *Let* $(\mathsf{X},\mathscr{A})$ *be a measurable space, let* $\mathsf{f}\colon\mathsf{X}\to\overline{\mathbb{R}}$ *be* $\mathscr{A}$*-measurable, and let* $\mathsf{A}\in\mathscr{A}$*. Then* $\mathsf{f}|\mathsf{A}$ *is* $\mathscr{A}_{\mathsf{A}}$*-measurable.*

Moreover, if $\mathsf{B}=\mathsf{X}\setminus\mathsf{A}$ *and if we have* $\mathscr{A}_{\mathsf{A}}$*- and* $\mathscr{A}_{\mathsf{B}}$*-measurable functions* $\mathsf{f}_{\mathsf{A}}\colon\mathsf{A}\to\overline{\mathbb{R}}$ *and* $\mathsf{f}_{\mathsf{B}}\colon\mathsf{B}\to\overline{\mathbb{R}}$*, respectively, then the function* $\mathsf{f}\colon\mathsf{X}\to\overline{\mathbb{R}}$ *defined by*

$$\mathsf{f}(\mathsf{x})=\begin{cases}\mathsf{f}_{\mathsf{A}}(\mathsf{x}), & \mathsf{x}\in\mathsf{A},\\ \mathsf{f}_{\mathsf{B}}(\mathsf{x}), & \mathsf{x}\in\mathsf{B}\end{cases}$$

is $\mathscr{A}$*-measurable.*

Proof Let $E\in\mathscr{A}$ so that $A\cap E\in\mathscr{A}_A$. Then, by Proposition 1.3.5,

$$f^{-1}(A\cap E)=f^{-1}(A)\cap f^{-1}(E),$$

and from this we deduce that $f^{-1}(A\cap E)$ is the intersection of measurable sets, and so measurable.

For the second assertion of the proposition, let $E\in\mathscr{A}$ and write $E=(A\cap E)\cup(B\cap E)$. Then, again by Proposition 1.3.5,

$$f^{-1}(E)=f^{-1}(A\cap E)\cup f^{-1}(B\cap E)=f_A^{-1}(A\cap E)\cup f_B^{-1}(B\cap E).$$

Since f_A and f_B are $\mathscr{A}_A$- and $\mathscr{A}_B$-measurable, $f_A^{-1}(A\cap E)\in\mathscr{A}_A$ and $f_B^{-1}(B\cap E)\in\mathscr{A}_B$. Since $\mathscr{A}_A,\mathscr{A}_B\subseteq\mathscr{A}$, $f^{-1}(E)$ is the union of $\mathscr{A}$-measurable sets, and so is $\mathscr{A}$-measurable. ■

Let us consider the rôle of measurability with respect to the operations of min and max.

5.6.16 Proposition (Measurability and max and min) *If* $(X, \mathscr{A})$ *is a measure space and if* $f, g \in L^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$*, then the functions*

$$X \ni x \mapsto \min\{f(x), g(x)\} \in \overline{\mathbb{R}}, \qquad X \ni x \mapsto \max\{f(x), g(x)\} \in \overline{\mathbb{R}}.$$

are $\mathscr{A}$*-measurable.*

Proof Let $a \in \mathbb{R}$ and note that

$$\{x \in X \mid \min\{f(x), g(x)\} \leq a\} = \{x \in X \mid f(x) \leq a\} \cup \{x \in X \mid g(x) \leq a\}$$

and

$$\{x \in X \mid \max\{f(x), g(x)\} \leq a\} = \{x \in X \mid f(x) \leq a\} \cap \{x \in X \mid g(x) \leq b\}.$$

Thus $\{x \in X \mid \min\{f(x), g(x)\} \leq a\}$ and $\{x \in X \mid \max\{f(x), g(x)\} \leq a\}$ are measurable and this gives the result. ■

The previous result has the following obvious corollary that will be useful when we define the integral.

5.6.17 Corollary (Measurability of positive and negative parts of a function) *Let* $(X, \mathscr{A})$ *be a measurable space and let* $f \in L^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$ *be* $\mathscr{A}$*-measurable. Then the functions* $f_-, f_+ : X \to \overline{\mathbb{R}}$ *defined by*

$$f_+(x) = \max\{f(x), 0\}, \qquad f_-(x) = \max\{-f(x), 0\}$$

are $\mathscr{A}$*-measurable.*

Proof By Example 5.6.8–1 the function $x \mapsto 0$ is $\mathscr{A}$-measurable. The corollary now follows immediately from Proposition 5.6.16. ■

5.6.3 Sequences of measurable functions

In Sections 3.5 and **??*missing stuff*** we considered sequences of continuous functions. We saw that notions of uniform convergence are important for the preservation of continuity of the limit function. Measurable functions are far more flexible in this regard, and so we are able to assert the measurability of a fairly general collection of operations applied to sequences of measurable functions. First let us define some notation to facilitate the statement of the result. We let $(X, \mathscr{A})$ be a measurable space with $S = (f_j)_{j \in \mathbb{Z}_{>0}}$ a sequence of $\mathscr{A}$-measurable functions. We then define functions $\inf S, \sup S, \liminf S, \limsup S : X \to \overline{\mathbb{R}}$ by

$$\begin{aligned} \inf S(x) &= \inf\{f_j(x) \mid j \in \mathbb{Z}_{>0}\}, & \sup S(x) &= \sup\{f_j(x) \mid j \in \mathbb{Z}_{>0}\}, \\ \liminf S(x) &= \liminf_{j \to \infty} f_j(x), & \limsup S(x) &= \limsup_{j \to \infty} f_j(x). \end{aligned}$$

Note that these four functions are always defined, regardless of the sequence. Let us also define

$$A_S = \{x \in X \mid \liminf S(x) = \limsup S(x)\}$$

and define $\lim S : A_S \to \overline{\mathbb{R}}$ by $\lim S(x) = \lim_{j \to \infty} f_j(x)$, noting that this is also a well-defined function. With this notation we have the following result.

5.6.18 Proposition (Limit operations on measurable functions) *Let $(\mathsf{X}, \mathscr{A})$ be a measurable space and let $\mathsf{S} = (\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$. Then the following statements hold:*

(i) $\inf \mathsf{S} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*;*

(ii) $\sup \mathsf{S} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*;*

(iii) $\liminf \mathsf{S} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*;*

(iv) $\limsup \mathsf{S} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*;*

(v) $\mathsf{A}_\mathsf{S} \in \mathscr{A}$ *and* $\lim \mathsf{S} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*.*

Proof (i) For $b \in \mathbb{R}$ we have

$$\{x \in X \mid \ \inf S(x) < b\} = \bigcup_{j\in\mathbb{Z}_{>0}} \{x \in X \mid \ f_j(x) < b\}.$$

Since the sets of the right are measurable, so is their union.

(ii) For $b \in \mathbb{R}$ we have

$$\{x \in X \mid \ \sup S(x) \le b\} = \bigcap_{j\in\mathbb{Z}_{>0}} \{x \in X \mid \ f_j(x) \le b\}.$$

Since the sets on the right are measurable, so is their intersection.

(iii) Define a sequence of functions $(\underline{f}_j)_{j\in\mathbb{Z}_{>0}}$ by $\underline{f}_j(x) = \sup_{k\ge j} f_k(x)$. These functions are $\mathscr{A}$-measurable by part (i). By Proposition 2.3.16 we have

$$\liminf_{j\to\infty} f_j(x) = \sup\left\{\underline{f}_k(x) \;\middle|\; k \in \mathbb{Z}_{>0}\right\},$$

and so this part of the result follows from part (i).

(iv) Define a sequence of functions $(\overline{f}_j)_{j\in\mathbb{Z}_{>0}}$ by $\overline{f}_j(x) = \sup_{k\ge j} f_k(x)$. These functions are $\mathscr{A}$-measurable by part (ii). By Proposition 2.3.15 we have

$$\limsup_{j\to\infty} f_j(x) = \inf\left\{\overline{f}_k(x) \;\middle|\; k \in \mathbb{Z}_{>0}\right\},$$

and so this part of the result follows from part (ii).

(v) Measurability of A_S follows from parts (iii) and (iv), along with Lemma 1 from the proof of Proposition 5.6.11. Now let $b \in \mathbb{R}$ and note that

$$\{x \in A_S \mid \ \lim f(x) \le b\} = A_S \cap \{x \in X \mid \ \limsup S(x) \le b\}.$$

The set on the right is the intersection of measurable sets and so is measurable. This then gives $\mathscr{A}$-measurability of $\lim S$ by Proposition 5.2.6. ■

The following corollary will come up often. Note that this result is unlike most of the results thus far in this section in that it depends on a measure.

5.6.19 Corollary (Measurability of almost everywhere convergent sequences) *Let* $(\mathrm{X}, \mathscr{A}, \mu)$ *be a complete measure space, let* $(\mathrm{f}_j)_{j\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \overline{\mathbb{R}})$*, and let* $\mathrm{f}\colon \mathrm{X} \to \overline{\mathbb{R}}$ *be such that*

$$\mu\Big(\mathrm{X} \setminus \Big\{\mathrm{x} \in \mathrm{X} \;\Big|\; \mathrm{f}(\mathrm{x}) = \lim_{j\to\infty} \mathrm{f}_j(\mathrm{x})\Big\}\Big) = 0.$$

Then f *is* $\mathscr{A}$*-measurable.*

Proof Let $S = (f_j)_{j\in\mathbb{Z}_{>0}}$ and define

$$B_S = \Big\{x \in X \;\Big|\; f(x) = \lim_{j\to\infty} f_j(x)\Big\}.$$

From Proposition 5.6.18 the function $\liminf S$ is $\mathscr{A}$-measurable. Since f and $\liminf S$ agree except on the set B_S which has measure zero, it follows from Proposition 5.6.10 that f is $\mathscr{A}$-measurable since μ is complete. ■

The preceding few results had to do with the measurability of various sorts of limits of measurable functions. Let us now study systematically the various sorts of convergence that may be experienced by sequences of measurable functions. In ***missing stuff*** we described the notions of pointwise and uniform convergence in a general way using topological ideas. These definitions carry over to sequences of functions defined on measure spaces, but there are additional notions arising from the measure theoretic setting, as the following definitions make clear.

5.6.20 Definition (Modes of convergence for sequences of measurable functions) Let $(X, \mathscr{A}, \mu)$ be a measure space, let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$, and let $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$. The sequence

(i) ***converges pointwise*** to f if $\lim_{j\to\infty} f_j(x) = f(x)$ for every $x \in X$,

(ii) ***converges pointwise almost everywhere*** to f if

$$\mu\Big(X \setminus \Big\{x \in X \;\Big|\; f(x) = \lim_{j\to\infty} f_j(x)\Big\}\Big) = 0,$$

(iii) ***converges uniformly*** to f if, for every $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \epsilon$ for every $x \in X$ and for every $j \geq N$,

(iv) ***converges almost uniformly*** to f if, for every $\delta \in \mathbb{R}_{>0}$, there exists a set $E_\delta \subseteq X$ having the following properties:

(a) $\mu(E_\delta) < \delta$;

(b) for every $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \epsilon$ for every $x \in X \setminus E_\delta$ and for every $j \geq N$,

and

(v) ***converges in measure*** to f if, for every $\epsilon \in \mathbb{R}_{>0}$,

$$\lim_{j\to\infty} \mu\big(\{x \in X \mid |f(x) - f_j(x)| > \epsilon\}\big) = 0.$$

•

Some of the relationships between the various notions of convergence are obvious. For example, the implications

$$\text{(iv)} \Longleftarrow \text{(iii)} \implies \text{(i)} \implies \text{(ii)}$$

obviously hold. Moreover, the converse implications of some of the preceding implications fairly obviously do not hold in general. For example, we know from Section 3.5.2 that generally (i)$\not\Rightarrow$(iii). It is also pretty evident that generally (ii)$\not\Rightarrow$(i); see Exercise 5.6.4. Let us now explore the possibility of other implications. The first result shows that, perhaps a little surprisingly, (ii) implies (iv) when the functions in the sequence and the limit function are $\mathbb{R}$-valued, and when the measure is finite.

5.6.21 Theorem (Egorov's[7] Theorem) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a finite measure space and let* $\mathsf{f}_j, \mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$, $j \in \mathbb{Z}_{>0}$, *have the following properties:*

(i) the sets $\{x \in \mathsf{X} \mid \mathsf{f}_j(x) \notin \mathbb{R}\}$, $j \in \mathbb{Z}_{>0}$, *and* $\{x \in \mathsf{X} \mid \mathsf{f}(x) \notin \mathbb{R}\}$ *have measure zero;*

(ii) $(\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ *converges pointwise almost everywhere to* f.

Then $(\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ *converges almost uniformly to* f.

Proof First let us suppose that f and f_j, $j \in \mathbb{Z}_{>0}$, are $\mathbb{R}$-valued. For $k, m \in \mathbb{Z}_{>0}$ define

$$E_{km} = \left\{x \in X \;\middle|\; |f(x) - f_m(x)| < \tfrac{1}{k}\right\}.$$

Since $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges almost everywhere to f, there exists a set $Z \subseteq X$ such that

1. $\mu(Z) = 0$ and
2. for $k \in \mathbb{Z}_{>0}$ and $x \in X \setminus Z$, there exists $m \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \frac{1}{k}$ for $j \geq m$.

That is to say,

$$X \setminus Z \subseteq \bigcup_{n\in\mathbb{Z}_{>0}} \bigcap_{m\geq n} E_{km} \quad \implies \quad Z \subseteq \bigcap_{n\in\mathbb{Z}_{>0}} \bigcup_{m\geq n} X \setminus E_{km}$$

for every $k \in \mathbb{Z}_{>0}$, using De Morgan's Laws. Denote $A_{kn} = \cup_{m\geq n} X \setminus E_{km}$. Note that $A_{kn} \supseteq A_{k(n+1)}$ for every $k, n \in \mathbb{Z}_{>0}$, and that $\cap_{n\in\mathbb{Z}_{>0}} A_{kn} \subseteq Z$ which implies that $\cap_{n\in\mathbb{Z}_{>0}} A_{kn}$ has zero measure, being a subset of a set with zero measure. Let $Z_k = \cap_{n\in\mathbb{Z}_{>0}} A_{kn}$ and note that

$$\lim_{n\to\infty} \mu(A_{kn}) = \lim_{n\to\infty} \mu(A_{kn} \setminus Z_k) = 0$$

using Proposition 5.3.3.

Let $\delta \in \mathbb{R}_{>0}$. For $k \in \mathbb{Z}_{>0}$ let $N_k \in \mathbb{Z}_{>0}$ be such that $\mu(A_{kn}) < \frac{\delta}{2^k}$ for $n \geq N_k$. Define

$$E_\delta = \bigcup_{k\in\mathbb{Z}_{>0}} A_{kN_k}.$$

Then

$$\mu(E_\delta) \leq \sum_{k=1}^{\infty} \mu(A_{kN_k}) < \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \delta$$

[7] Dimitri Fedorovich Egorov (1869–1931) Was a Russian mathematician whose main mathematical contributions were to differential geometry and analysis.

by Example 2.4.2–??.

Now let $\epsilon \in \mathbb{R}_{>0}$ and take $K \in \mathbb{Z}_{>0}$ such that $\frac{1}{K} < \epsilon$. If $x \in X \setminus E_\delta$ we have, by definition of E_δ and De Morgan's Laws,

$$x \in \bigcap_{k \in \mathbb{Z}_{>0}} \bigcap_{m \geq N_k} E_{km},$$

which implies in particular that $x \in E_{Km}$ whenever $m \geq N_K$. That is to say, if $j \geq N_K$ then $|f(x) - f_j(x)| < \epsilon$ for every $x \in X \setminus E_\delta$, as desired.

To conclude the proof, let us relax the assumption made above that f and f_j, $j \in \mathbb{Z}_{>0}$, are $\mathbb{R}$-valued. Define

$$N = \{x \in X \mid f(x) \notin \mathbb{R}\}, \quad N_j = \{x \in X \mid f_j(x) \notin \mathbb{R}\}, \qquad j \in \mathbb{Z}_{>0}.$$

If $Z = N \cup (\cup_{j \in \mathbb{Z}_{>0}} N_j)$ then Z is a measurable set with zero measure, being a countable union of sets with zero measure. The hypotheses from the first part of the proof hold for $X \setminus Z$ and for f and f_j, $j \in \mathbb{Z}_{>0}$, restricted to $X \setminus Z$. That is to say, for every $\delta \in \mathbb{R}_{>0}$ there exists a set $E'_\delta \subseteq (X \setminus Z)$ such that $\mu(E'_\delta) < \delta$ and such that, for every $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_j(x)| < \epsilon$ for every $x \in X \setminus (Z \cup E'_\delta)$ and for every $j \geq N$. Now let $\delta \in \mathbb{R}_{>0}$ and take $E_\delta = E'_\delta \cup Z$. Note that $\mu(E_\delta) = \mu(E'_\delta) < \delta$. Now, for $\epsilon \in \mathbb{R}_{>0}$ let N be chosen as above, so that $|f(x) - f_j(x)| < \epsilon$ for every $x \in X \setminus E_\delta$ and for every $j \geq N$. This gives almost uniform convergence of $(f_j)_{j \in \mathbb{Z}_{>0}}$ to f, as desired. ■

Note that the theorem allows us to immediately conclude that generally (iv)$\not\Rightarrow$(iii) from Definition 5.6.20. Indeed, suppose that $(f_j)_{j \in \mathbb{Z}_{>0}}$ is a sequence of $\mathbb{R}$-valued functions on $[0, 1]$ that converges pointwise, but not uniformly, to a function f. Then the preceding theorem implies that the sequence converges almost uniformly to f.

The next example shows that finiteness of the measure space in Egorov's Theorem is necessary.

5.6.22 Example (Egorov's Theorem generally fails for measure spaces that are not finite) We take $X = \mathbb{R}$, $\mathscr{A} = \mathscr{L}(\mathbb{R})$, and $\mu = \lambda$. We consider the sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathsf{L}^{(0)}((\mathbb{R}, \mathscr{L}(\mathbb{R})); \mathbb{R})$ defined by $f_j = \chi_{[j, j+1)}$. We also define $f \in \mathsf{L}^{(0)}((\mathbb{R}, \mathscr{L}(\mathbb{R})); \mathbb{R})$ by $f(x) = 0$ for all $x \in \mathbb{R}$. We claim that $(f_j)_{j \in \mathbb{Z}_{>0}}$ converges pointwise, and so pointwise almost everywhere, to f, but does not converge almost uniformly to f. To verify pointwise convergence, let $x \in \mathbb{R}$ and choose $N \in \mathbb{Z}_{>0}$ such that $N + 1 > x$, Then we have $f_j(x) = 0$ for all $j \geq N$, so verifying pointwise convergence to f. To see that the sequence does not converge almost uniformly, let $\delta, \epsilon \in (0, 1)$ and suppose that $E_\delta \subseteq \mathbb{R}$ is such that there exists $N \in \mathbb{Z}_{>0}$ for which $|f(x) - f_j(x)| < \epsilon$ for $x \in \mathbb{R} \setminus E_\delta$ and for $j \geq N$. This means that $|f_j(x)| \geq \epsilon$ on a set A contained in E_δ for $j \geq N$. But this implies that $[N, N + 1) \subseteq E_\delta$, implying that $\mu(E_\delta) > \delta$. This precludes almost uniform convergence. •

The preceding discussion concerning the relationships between modes of convergence has not involved convergence in measure. Let us now investigate the rôle of convergence in measure relative to the other modes of convergence. The first result establishes that for finite measure spaces we have the implication (ii) $\Longrightarrow$ (v) from Definition 5.6.20.

5.6.23 Proposition (Almost everywhere pointwise convergence sometimes implies convergence in measure) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a finite measure space. Consider a sequence* $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *and a function* f *in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *with the following properties:*

(i) the sets $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) \notin \mathbb{R}\}$ *and* $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}_\mathsf{j}(\mathsf{x}) \notin \mathbb{R}\}$, $\mathsf{j} \in \mathbb{Z}_{>0}$, *have measure zero;*

(ii) $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *converges pointwise almost everywhere to* f.

Then $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *converges in measure to* f.

Proof First suppose that f and f_j, $j \in \mathbb{Z}_{>0}$, are $\mathbb{R}$-valued. Let $\epsilon \in \mathbb{R}_{>0}$ and define

$$A_{\epsilon,j} = \{x \in X \mid |f(x) - f_j(x)| > \epsilon\}, \qquad j \in \mathbb{Z}_{>0}$$

and $B_{\epsilon,k} = \cup_{j=1}^k A_{\epsilon,j}$, $k \in \mathbb{Z}_{>0}$. Then we have $B_{k+1} \supseteq B_k$ for $k \in \mathbb{Z}_{>0}$ and

$$\cap_{k\in\mathbb{Z}_{>0}} B_{\epsilon,k} \subseteq \{x \in X \mid (f_j(x))_{j\in\mathbb{Z}_{>0}} \text{ does not converge to } f(x)\}$$

Therefore, $\mu(\cap_{k\in\mathbb{Z}_{>0}} B_{\epsilon,k}) = 0$ and so, by Proposition 5.3.3, $\lim_{k\to\infty} B_{\epsilon,k} = 0$. Therefore, since $A_{\epsilon,j} \subseteq B_{\epsilon,j}$ for $j \in \mathbb{Z}_{>0}$, we have $\lim_{j\to\infty} A_{\epsilon,j} = 0$. This is exactly the statement that $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to f in measure.

To complete the proof, suppose that f_j, $j \in \mathbb{Z}_{>0}$, are not necessarily $\mathbb{R}$-valued. Let

$$N = \{x \in X \mid f(x) \notin \mathbb{R}\}, \quad N_j = \{x \in X \mid f_j(x) \notin \mathbb{R}\}, \qquad j \in \mathbb{Z}_{>0}$$

so that $Z = N \cup (\cup_{j\in\mathbb{Z}_{>0}} N_j)$ is a measurable set with zero measure, it being a countable union of sets with zero measure. The first part of the proof then applies for $X \setminus Z$ and for f and f_j, $j \in \mathbb{Z}_{>0}$, restricted to $X \setminus Z$. Thus, for $\epsilon \in \mathbb{R}_{>0}$ we have

$$\lim_{j\to\infty} \mu(\{x \in X \setminus Z \mid |f(x) - f_j(x)| > \epsilon\}) = 0.$$

Since

$$\begin{aligned}\{x \in X \mid |f(x) - f_j(x)| > \epsilon\} &= \{x \in X \setminus Z \mid |f(x) - f_j(x)| > \epsilon\} \cup \{x \in Z \mid |f(x) - f_j(x)| > \epsilon\} \\ &\subseteq \{x \in X \setminus Z \mid |f(x) - f_j(x)| > \epsilon\} \cup Z,\end{aligned}$$

we have

$$\lim_{j\to\infty} \mu(\{x \in X \mid |f(x) - f_j(x)| > \epsilon\}) = 0,$$

giving convergence in measure as desired. ■

The condition that the measure space be finite is generally necessary in the preceding result.

5.6.24 Example (Almost everywhere pointwise convergence does not always imply convergence in measure) Here we take $X = \mathbb{R}$, $\mathscr{A} = \mathscr{L}(\mathbb{R})$, and $\mu = \lambda$. We define a sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ and a function f in $\mathsf{L}^{(0)}((\mathbb{R}, \mathscr{L}(\mathbb{R})); \mathbb{R})$ by $f_j = \chi_{[j,j+1)}$ and $f(x) = 0$ for $x \in \mathbb{R}$. We saw in Example 5.6.22 that the sequence converges pointwise to f, and so converges pointwise almost everywhere to f. However, if $\epsilon \in (0, 1)$ then

$$\lambda(\{x \in \mathbb{R} \mid |f(x) - f_j(x)| > \epsilon\}) = 1$$

which clearly precludes the sequence from converging to f in measure. •

Now let us investigate the extent to which convergence in measure implies almost everywhere pointwise convergence. The following example shows that the general implication fails to hold, even for finite measure spaces.

5.6.25 Example (Convergence in measure does not imply almost everywhere pointwise convergence) We take $X = [0,1)$, $\mathscr{A} = \mathscr{L}([0,1))$, and $\mu = \lambda_{[0,1)}$. We define a sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{L}^{(0)}(([0,1),\mathscr{L}([0,1)));\mathbb{R})$ as follows. For $k \in \mathbb{Z}_{\geq 0}$ we define $f_{2^k}, f_{2^k+1}, \dots, f_{2^{k+1}-1}$ by $f_{2^k+j} = \chi_{[j2^{-k},(j+1)2^{-k})}$, $j \in \{0,1,\dots,2^k-1\}$. Thus, for example,

$$\begin{aligned}
&f_1 = \chi_{[0,1)},\\
&f_2 = \chi_{[0,\frac{1}{2})},\ f_3 = \chi_{[\frac{1}{2},1)},\\
&f_4 = \chi_{[0,\frac{1}{4})},\ f_5 = \chi_{[\frac{1}{4},\frac{1}{2})},\ f_6 = \chi_{[\frac{1}{2},\frac{3}{4})},\ f_7 = \chi_{[\frac{3}{4},1)}.
\end{aligned}$$

We also define $f \in \mathsf{L}^{(0)}(([0,1),\mathscr{L}([0,1)));\mathbb{R})$ by $f(x) = 0$ for $x \in [0,1)$. We claim that this sequence converges in measure to f, but does not converge pointwise almost everywhere to f.

To verify convergence in measure, let $\epsilon \in \mathbb{R}_{>0}$ and note that for any $j \in \mathbb{Z}_{>0}$ we have

$$\{x \in [0,1) \mid |f(x) - f_j(x)| > \epsilon\} \subseteq \{x \in [0,1) \mid |f_j(x)| > 0\}.$$

If $j \in \{2^k, 2^k+1, \dots, 2^{k+1}-1\}$ then

$$\lambda(\{x \in [0,1) \mid |f_j(x)| > 0\}) = 2^{-k}.$$

Therefore, it follows that

$$\lim_{j\to\infty} \lambda(\{x \in [0,1) \mid |f_j(x)| > 0\}) = 0,$$

giving convergence in measure.

Now we verify that the sequence does not converge pointwise almost everywhere. Let $x \in [0,1)$ and let $N \in \mathbb{Z}_{>0}$. Choose $k \in \mathbb{Z}_{>0}$ such that $2^k > N$ and choose $j \in \{0,1,\dots,2^k-1\}$ such that $x \in [j2^{-k},(j+1)2^{-k})$. Then, for $m \in \{2^k, 2^k+1, \dots, 2^{k+1}-1\}$ we have

$$f_m(x) = \begin{cases} 1, & m = 2^k + j,\\ 0, & \text{otherwise.}\end{cases}$$

Thus, no matter how large we choose N, there are terms beyond the Nth term in the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ that have value 1 and terms beyond the Nth term in the sequence $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ that have value 0. This precludes pointwise convergence at x. Since this is true for every $x \in [0,1)$ it follows that almost everywhere pointwise convergence is precluded. Indeed, the sequence converges pointwise nowhere. •

The situation is not entirely hopeless, however. Indeed, one has the following result.

5.6.26 Proposition (Convergence in measure implies almost everywhere pointwise convergence of a subsequence) *Let $(\mathsf{X},\mathscr{A},\mu)$ be a measure space, let $(\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}})$ and let $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}})$ satisfy the following:*

(i) the sets $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) \notin \mathbb{R}\}$ and $\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}_j(\mathsf{x}) \notin \mathbb{R}\}$ have measure zero;

(ii) the sequence $(\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ converges to f in measure.

Then there exists a subsequence of $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *which converges pointwise almost everywhere to* f.

Proof Define a strictly increasing sequence $(j_k)_{k\in\mathbb{Z}_{>0}}$ in $\mathbb{Z}_{>0}$ as follows. Let j_1 be such that

$$\mu(\{x \in X \mid |f(x) - f_{j_1}(x)| > 1\}) \leq \tfrac{1}{2},$$

this being possible by definition of convergence in measure. Then suppose that $j_1, \dots, j_k$ have been defined. Define j_{k+1} such that $j_{k+1} > j_k$ and such that

$$\mu(\{x \in X \mid |f(x) - f_{j_{k+1}}(x)| > \tfrac{1}{k+1}\} \leq \tfrac{1}{2^{k+1}},$$

this again being possible by definition of convergence in measure. Now define

$$A_k = \{x \in X \mid |f(x) - f_{j_k}(x)| < \tfrac{1}{k}\}, \qquad k \in \mathbb{Z}_{>0},$$

and $B_m = \cup_{k=m}^{\infty} A_k$, $m \in \mathbb{Z}_{>0}$. Note that $B_{m+1} \supseteq B_m$ for $m \in \mathbb{Z}_{>0}$. Moreover,

$$\mu(B_m) \leq \sum_{k=m}^{\infty} \mu(A_k) \leq \sum_{k=m}^{\infty} \frac{1}{2^k} = \frac{1}{2^{m-1}} \sum_{k=1}^{\infty} \frac{1}{2^k} = \frac{1}{2^{m-1}}$$

by Example 2.4.2–**??**. Therefore, by Proposition 5.3.3,

$$\mu(\cap_{m=1}^{\infty} B_m) = \lim_{m\to\infty} \mu(B_m) \leq \lim_{m\to\infty} \frac{1}{2^{m-1}} = 0.$$

Now, if $x \notin \cap_{m=1}^{\infty} B_m$ there exists $m \in \mathbb{Z}_{>0}$ such that $x \notin B_m$. Thus $x \notin \cup_{k=m}^{\infty} A_k$ and so

$$|f(x) - f_{j_k}(x)| < \tfrac{1}{2^k}, \qquad k \geq m.$$

Thus $\lim_{k\to\infty} f_{j_k}(x) = f(x)$. Thus $(f_{j_k})_{k\in\mathbb{Z}_{>0}}$ converges pointwise to f on $X \setminus (\cap_{m=1}^{\infty} B_m)$. This gives almost everywhere pointwise convergence of this subsequence to f. ■

5.6.4 ℂ- and vector-valued measurable functions

It is important to be able to talk about functions taking values in spaces more interesting than $\overline{\mathbb{R}}$. In particular, ℂ-valued functions will be frequently encountered in these volumes. Here we allow this by considering functions taking values in $\mathbb{R}^n$.

First let us define what we mean by a measurable $\mathbb{R}^n$-valued function.

5.6.27 Definition (Measurable vector-valued function) For a measurable space $(X, \mathscr{A})$, a function $f\colon X \to \mathbb{R}^n$ is ***$\mathscr{A}$-measurable*** if its components $f_1, \dots, f_n\colon X \to \mathbb{R}$ are measurable in the sense of Definition 5.6.5. We denote the set of measurable $\mathbb{R}^n$-valued maps by $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R}^n)$, or simply by $\mathsf{L}^{(0)}(X; \mathbb{R}^n)$ with the understanding that the σ-algebra $\mathscr{A}$ is implicit. •

Let us relate this notion of measurability to that in Definition 5.6.1.

5.6.28 Proposition (Characterisation of vector-valued measurable functions) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and for a function* $\mathbf{f}\colon \mathsf{X} \to \mathbb{R}^{\mathsf{n}}$ *the following statements are equivalent:*

(i) $\mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}})$*;*

(ii) $\mathbf{f}$ *is* $(\mathscr{A}, \mathscr{B}(\mathbb{R}^{\mathsf{n}}))$*-measurable.*

Proof Suppose that $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R}^n)$. By Propositions 5.4.9 it follows that $f_j^{-1}((-\infty, b_j]) \in \mathscr{A}$ for every $b_j \in \mathbb{R}$ and for $j \in \{1, \dots, n\}$. Now note that

$$f^{-1}((-\infty, b_1] \times \cdots \times (-\infty, b_n]) = f_1^{-1}((-\infty, b_1]) \cap \cdots \cap f_n^{-1}((-\infty, b_n]) \in \mathscr{A}$$

for every $b_1, \dots, b_n \in \mathbb{R}$. By Propositions 5.5.9 and 5.6.2 it follows that f is $(\mathscr{A}, \mathscr{B}(\mathbb{R}^n))$-measurable.

Next suppose that f is $(\mathscr{A}, \mathscr{B}(\mathbb{R}^n))$-measurable. Then, for $j \in \{1, \dots, n\}$ and $b_j \in \mathbb{R}$,

$$\begin{aligned} f_j^{-1}((-\infty, b_j]) &= X \cap \cdots \cap f_j^{-1}((-\infty, b_j]) \cap \cdots \cap X \\ &= f_1^{-1}(\mathbb{R}) \cap \cdots \cap f_j^{-1}((-\infty, b_j]) \cap \cdots \cap f_n^{-1}(\mathbb{R}) \\ &= f^{-1}(\mathbb{R} \times \cdots \times (-\infty, b_j] \times \cdots \times \mathbb{R}). \end{aligned}$$

Since $\mathbb{R} \times \cdots \times (-\infty, b_j] \times \cdots \times \mathbb{R}$ is a Borel set (it is closed), it follows that $f_j^{-1}((-\infty, b_j])$ is a Borel set, and so the result follows from Propositions 5.4.9 and 5.6.2. ■

The definition of measurable $\mathbb{C}$-valued functions follows directly from the preceding constructions. Indeed, we note that $\mathbb{C}$ is isomorphic as a $\mathbb{R}$-vector space to $\mathbb{R}^n$ via the isomorphism $z \mapsto (\mathrm{Re}(z), \mathrm{Im}(z))$. Thus the following definition simply specialises the above general definition.

5.6.29 Definition (Measurable $\mathbb{C}$-valued functions) For a measurable space $(X, \mathscr{A})$, a function $f\colon X \to \mathbb{C}$ is ***$\mathscr{A}$-measurable*** if the $\mathbb{R}$-valued functions

$$\mathrm{Re}(f)\colon x \mapsto \mathrm{Re}(f(x)), \quad \mathrm{Im}(f)\colon x \mapsto \mathrm{Im}(f(x))$$

are measurable in the sense of Definition 5.6.5. We denote the set of measurable $\mathbb{C}$-valued maps by $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{C})$, with the understanding that the σ-algebra $\mathscr{A}$ is implicit. •

It is straightforward to adapt the results concerning operations on measurable functions in Section 5.6.2 to vector-valued functions. Let us record this here for $\mathbb{R}^n$-valued functions, noting that these results apply immediately to $\mathbb{C}$-valued functions.

5.6.30 Proposition (Algebraic operations on measurable functions) *Let* $(\mathsf{X}, \mathscr{A})$ *be a measurable space, let* $\mathbf{f}, \mathbf{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}})$*, and let* $\mathsf{a} \in \mathbb{R}$*. Then the functions* $\mathbf{f} + \mathbf{g}$ *and* $\mathsf{a}\mathbf{f}$ *defined by*

$$(\mathbf{f} + \mathbf{g})(\mathsf{x}) = \mathbf{f}(\mathsf{x}) + \mathbf{g}(\mathsf{x}), \qquad (\mathsf{a}\mathbf{f})(\mathsf{x}) = \mathsf{a}(\mathbf{f}(\mathsf{x}))$$

are $\mathscr{A}$*-measurable.*

Proof This follows directly from the definition of $\mathscr{A}$-measurable vector-valued functions, the definitions of vector addition and scalar multiplication, and Proposition 5.6.11. ■

Next we consider compositions of measurable functions with functions between Euclidean spaces.

5.6.31 Proposition (Composition and measurable functions) *Let* $(\mathsf{X}, \mathscr{A})$ *be a measurable space, let* $\mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}})$*, and let* $\boldsymbol{\phi}\colon \mathbb{R}^{\mathsf{n}} \to \mathbb{R}^{\mathsf{m}}$ *be* $\mathscr{B}(\overline{\mathbb{R}})$*-measurable. Then* $\boldsymbol{\phi} \circ \mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{m}})$*.*

Proof Let $B \in \mathscr{B}(\mathbb{R}^m)$. By assumption and by Proposition 5.6.6 we have $\boldsymbol{\phi}^{-1}(B) \in \mathscr{B}(\mathbb{R}^n)$. Thus, using Exercise 1.3.2,

$$(\boldsymbol{\phi} \circ \boldsymbol{f})^{-1}(B) = \boldsymbol{f}^{-1}(\boldsymbol{\phi}^{-1}(B)) \in \mathscr{A},$$

and so $\boldsymbol{\phi} \circ \boldsymbol{f}$ is $\mathscr{A}$-measurable, as desired. ■

5.6.32 Corollary (Composition by continuous functions and measurability) *Let* $(\mathsf{X}, \mathscr{A})$ *be a measurable space, let* $\mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}})$*, and let* $\boldsymbol{\phi}\colon \mathbb{R}^{\mathsf{n}} \to \mathbb{R}^{\mathsf{n}}$ *be continuous. Then* $\boldsymbol{\phi} \circ \mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{m}})$*.*

Proof This follows from Proposition 5.6.31, along with Example 5.6.3. ■

5.6.33 Corollary (Measurability of norms of functions) *Let* $(\mathsf{X}, \mathscr{A})$ *be a measurable space, let* $\mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}})$*, and let* $\boldsymbol{\phi}\colon \mathbb{R}^{\mathsf{n}} \to \mathbb{R}^{\mathsf{n}}$ *be continuous. Then the function* $\mathsf{x} \mapsto \|\mathbf{f}(\mathsf{x})\|_{\mathbb{R}^{\mathsf{n}}}$ *is* $\mathscr{A}$*-measurable.*

Proof This follows from the previous corollary, along with continuity of the norm (***missing stuff***). ■

Next we consider the restrictions of measurable functions, recalling from Proposition 5.2.6 the definition of the restriction $\mathscr{A}_A$ of a measurable space $(X, \mathscr{A})$ to a measurable subset $A \in \mathscr{A}$.

5.6.34 Proposition (Measurability and restriction) *Let* $(\mathsf{X}, \mathscr{A})$ *be a measurable space, let* $\mathbf{f}\colon \mathsf{X} \to \mathbb{R}^{\mathsf{n}}$ *be* $\mathscr{A}$*-measurable, and let* $\mathsf{A} \in \mathscr{A}$*. Then* $\mathbf{f}|\mathsf{A}$ *is* $\mathscr{A}_{\mathsf{A}}$*-measurable.*

Moreover, if $\mathsf{B} = \mathsf{X} \setminus \mathsf{A}$ *and if we have* $\mathscr{A}_{\mathsf{A}}$*- and* $\mathscr{A}_{\mathsf{B}}$*-measurable functions* $\mathbf{f}_{\mathsf{A}}\colon \mathsf{A} \to \mathbb{R}^{\mathsf{n}}$ *and* $\mathbf{f}_{\mathsf{B}}\colon \mathsf{B} \to \mathbb{R}^{\mathsf{n}}$*, respectively, then the function* $\mathbf{f}\colon \mathsf{X} \to \mathbb{R}^{\mathsf{n}}$ *defined by*

$$\mathbf{f}(\mathsf{x}) = \begin{cases} \mathbf{f}_{\mathsf{A}}(\mathsf{x}), & \mathsf{x} \in \mathsf{A}, \\ \mathbf{f}_{\mathsf{B}}(\mathsf{x}), & \mathsf{x} \in \mathsf{B} \end{cases}$$

is $\mathscr{A}$*-measurable.*

Proof Let $B \in \mathscr{A}$ so that $A \cap B \in \mathscr{A}_A$. Then, by Proposition 1.3.5,

$$\boldsymbol{f}^{-1}(A \cap B) = \boldsymbol{f}^{-1}(A) \cap \boldsymbol{f}^{-1}(B),$$

and from this we deduce that $\boldsymbol{f}^{-1}(A \cap B)$ is the intersection of measurable sets, and so measurable.

For the second assertion of the proposition, let $E \in \mathscr{A}$ and write $E = (A \cap E) \cup (B \cap E)$. Then, again by Proposition 1.3.5,

$$\boldsymbol{f}^{-1}(E) = \boldsymbol{f}^{-1}(A \cap E) \cup \boldsymbol{f}^{-1}(B \cap E) = \boldsymbol{f}_A^{-1}(A \cap E) \cup \boldsymbol{f}_B^{-1}(B \cap E).$$

Since $\boldsymbol{f}_A$ and $\boldsymbol{f}_B$ are $\mathscr{A}_A$- and $\mathscr{A}_B$-measurable, $\boldsymbol{f}_A^{-1}(A \cap E) \in \mathscr{A}_A$ and $\boldsymbol{f}_B^{-1}(B \cap E) \in \mathscr{A}_B$. Since $\mathscr{A}_A, \mathscr{A}_B \subseteq \mathscr{A}$, $\boldsymbol{f}^{-1}(E)$ is the union of $\mathscr{A}$-measurable sets, and so is $\mathscr{A}$-measurable. ■

Finally, we consider measurability of limits of vector-valued functions. We consider a sequence $S = (\boldsymbol{f}_j)_{j \in \mathbb{Z}_{>0}}$ of $\mathbb{R}^n$-valued functions on a measurable space $(X, \mathscr{A})$. Let us denote

$$A_S = \left\{x \in X \;\middle|\; \lim_{j \to \infty} \boldsymbol{f}_j(x) \text{ exists}\right\}$$

and define $\lim S \colon A_S \to \mathbb{R}^n$ by $\lim S(x) = \lim_{j \to \infty} \boldsymbol{f}_j(x)$. With this notation we have the following result.

5.6.35 Proposition (Pointwise limits of sequences of measurable functions) *Let* $(\mathsf{X}, \mathscr{A})$ *be a measurable space and let* $\mathsf{S} = (\mathbf{f}_\mathsf{j})_{\mathsf{j} \in \mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^\mathsf{n})$*. Then the set* A_S *and the function* $\lim \mathsf{S}$ *are* $\mathscr{A}$*-measurable.*

Proof Let $f_{1,j}, \dots, f_{n,j}$ be the components of $\boldsymbol{f}_j$, $j \in \mathbb{Z}_{>0}$, and for $k \in \{1, \dots, n\}$ define

$$A_{S,k} = \left\{x \in X \;\middle|\; \lim_{j \to \infty} f_{k,j}(x) \text{ exists}\right\}.$$

Note that $A_S = \cap_{k=1}^n A_{S,k}$ so that A_S is measurable, being a finite intersection of measurable sets. From Propositions 5.6.15 and 5.6.18 it follows that the function

$$A_S \ni x \mapsto \lim_{j \to \infty} f_{k,j}(x) \in \mathbb{R}$$

is $\mathscr{A}$-measurable. The definition of measurability of vector-valued functions now gives the result. ■

For almost everywhere pointwise convergent sequences, this gives the following result.

5.6.36 Corollary (Measurability of almost everywhere convergent sequences) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a complete measure space, let* $(\mathbf{f}_\mathsf{j})_{\mathsf{j} \in \mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^\mathsf{n})$*, and let* $\mathbf{f} \colon \mathsf{X} \to \mathbb{R}^\mathsf{n}$ *be such that*

$$\mu\left(\mathsf{X} \setminus \left\{\mathsf{x} \in \mathsf{X} \;\middle|\; \mathbf{f}(\mathsf{x}) = \lim_{\mathsf{j} \to \infty} \mathbf{f}_\mathsf{j}(\mathsf{x})\right\}\right) = 0.$$

Then $\mathbf{f}$ *is* $\mathscr{A}$*-measurable.*

Proof Let $f_{1,j}, \dots, f_{n,j}$ be the components of $\boldsymbol{f}_j$, $j \in \mathbb{Z}_{>0}$, let $f_1, \dots, f_n$ be the components of $\boldsymbol{f}$, and define

$$B_k = \left\{x \in A \;\middle|\; f_k(x) = \lim_{j \to \infty} f_{k,j}(x)\right\}, \qquad k \in \{1, \dots, n\},$$

and

$$B = \left\{x \in A \;\middle|\; f(x) = \lim_{j \to \infty} f_j(x)\right\}.$$

Note that $B = \cap_{k=1}^{n} B_k$. By hypothesis, $\mu(X \setminus B) = 0$, and we claim that $\mu(X \setminus B_k) = 0$ for $k \in \{1, \dots, n\}$. Indeed, suppose that $\mu(X \setminus B_k) > 0$ for some $k_0 \in \{1, \dots, n\}$. Then

$$\mu(X \setminus B) = \mu(X \setminus \cap_{j=1}^{k} B_k) = \mu(\cup_{k=1}^{n} X \setminus B_k) \geq \mu(X \setminus B_{k_0}) > 0,$$

contrary to our hypothesis. Since $\mu(X \setminus B_k) = 0$ for every $k \in \{1, \dots, n\}$, it follows from Corollary 5.6.19 that f_k is measurable, and so f is also measurable. ■

5.6.5 Simple functions and approximations of measurable functions

In this section we consider a specific class of measurable functions that will be fundamental to our construction of the integral in Section 5.7. There are various ways to characterise this class of functions, and the following result gives some of these.

5.6.37 Proposition (Characterisations of simple functions) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and a function* $\mathsf{f}\colon \mathsf{X} \to \overline{\mathbb{R}}$ *the following statements are equivalent:*

(i) $\mathrm{image}(\mathsf{f}) = \{\mathsf{a}_1, \dots, \mathsf{a}_k\} \subseteq \overline{\mathbb{R}}$ *and the sets* $\mathsf{f}^{-1}(\mathsf{a}_j)$, $\mathsf{j} \in \{1, \dots, \mathsf{k}\}$*, are measurable;*

(ii) there exists $\mathsf{B}_1, \dots, \mathsf{B}_m \in \mathscr{A}$ *and* $\mathsf{b}_1, \dots, \mathsf{b}_m \in \overline{\mathbb{R}}$ *such that* $\mathsf{f} = \sum_{j=1}^{m} \mathsf{b}_j \chi_{\mathsf{B}_j}$*;*

(iii) there exist pairwise disjoint sets $\mathsf{C}_1, \dots, \mathsf{C}_r \in \mathscr{A}$ *and* $\mathsf{c}_1, \dots, \mathsf{c}_r \in \overline{\mathbb{R}}$ *such that* $\mathsf{f} = \sum_{j=1}^{r} \mathsf{c}_j \chi_{\mathsf{C}_j}$.

Proof (i) $\Longrightarrow$ (ii) Given $f\colon X \to \overline{\mathbb{R}}$ satisfying condition (i), take $m = k$, $b_j = a_j$, and $B_j = f^{-1}(a_j)$, $j \in \{1, \dots, k\}$. If $x \in B_j$ then we clearly have $f(x) = b_j$ and so $f = \sum_{j=1}^{m} b_j \chi_{B_j}$, as desired.

(ii) $\Longrightarrow$ (iii) Let $f\colon X \to \overline{\mathbb{R}}$ satisfy condition (ii). If $x \in \cup_{j=1}^{m} B_j$ then there exists unique $j_1(x), \dots, j_{r(x)}(x) \in \{1, \dots, m\}$ such that $x \in C(x) \triangleq B_{j_1(x)} \cap \dots \cap B_{j_{r(x)}}(x)$, but $x \notin B_j$ for $j \notin \{j_1(x), \dots, j_{r(x)}(x)\}$. Moreover, $f(x) = b_{j_1(x)} + \dots + b_{j_{r(x)}(x)}$. Since there is a finite number of sets $B_1, \dots, B_m$ there are only finitely many possible intersections of these sets. Thus $\{C(x)\}_{x \in X} = \{C_1, \dots, C_r\}$ for disjoint sets $C_1, \dots, C_r$. Since each of the sets $C_1, \dots, C_r$ is a finite intersection of measurable sets, these sets are measurable. By construction of the sets $C(x)$, $x \in X$, the sets $C_1, \dots, C_m$ are pairwise disjoint. Moreover, the value of f on C_j is constant for $j \in \{1, \dots, r\}$. From these observations we immediately conclude that f satisfies property (iii).

(iii) $\Longrightarrow$ (i) First suppose that $\cup_{j=1}^{r} C_j = X$. Then, given $f\colon X \to \overline{\mathbb{R}}$ satisfying condition (iii), define $k = r$ and $a_j = c_j$, $j \in \{1, \dots, r\}$. Clearly $\mathrm{image}(f) = \{a_1, \dots, a_r\}$ and, since $f^{-1}(a_j)$ is a union of the measurable sets $C_1, \dots, C_r$ (it might be a union in case the numbers $c_1, \dots, c_r$ are not distinct), these sets are measurable. If $\cup_{j=1}^{r} C_j \subset X$ then define $k = r + 1$ and let $C_{r+1} = X \setminus \cup_{j=1}^{r} C_j$ and $c_{r+1} = 0$. The first part of the proof can now be repeated to give the desired conclusion in this case. ■

We now give a function having any of the preceding properties a name.

5.6.38 Definition (Simple function) If $(X, \mathscr{A})$ is a measurable space, a function $f\colon X \to \overline{\mathbb{R}}$ satisfying any one of the three equivalent properties of Proposition 5.6.37 is a ***simple function***. For any subset $I \subseteq \overline{\mathbb{R}}$ (typically we will be concerned with $I \in \{\mathbb{R}, \overline{\mathbb{R}}_{\geq 0}\}$) we denote

$$\mathsf{S}(X; I) = \{f\colon X \to I \mid f \text{ is simple}\},$$

with the understanding that the σ-algebra $\mathscr{A}$ is implicit. •

Simple functions can be thought of playing for the integral on measure spaces the rôle of step functions in the construction of the Riemann integral. For the Riemann integral, Riemann integrable functions are *defined* by their ability to be well approximated by step functions. For the integral defined on measure spaces, there exists a notion, definable only in terms of measurable sets, of a class of functions that are well approximated by simple functions. These are none other than the measurable functions that we have been talking about in this section. The following result illustrates this.

5.6.39 Proposition (Approximations of measurable functions by simple functions) *For a measurable space* $(\mathrm{X}, \mathscr{A})$ *and for an* $\mathscr{A}$*-measurable function* $\mathrm{f}\colon \mathrm{A} \to \overline{\mathbb{R}}$*, the following statements hold:*

(i) there exists a sequence $(\mathrm{f}_\mathrm{k})_{\mathrm{k}\in\mathbb{Z}_{>0}}$ *of simple functions having the property that, for each* $\mathrm{x} \in \mathrm{X}$*, we have*
$$\lim_{\mathrm{k}\to\infty} \mathrm{f}_\mathrm{k}(\mathrm{x}) = \mathrm{f}(\mathrm{x});$$

(ii) if f *is* $\overline{\mathbb{R}}_{\geq 0}$*-valued, the sequence* $(\mathrm{f}_\mathrm{k})_{\mathrm{k}\in\mathbb{Z}_{>0}}$ *of part (i) may be chosen so that the functions are* $\mathbb{R}_{\geq 0}$*-valued, and so that, for each* $\mathrm{x} \in \mathrm{X}$*, the sequence* $(\mathrm{f}_\mathrm{k}(\mathrm{x}))_{\mathrm{k}\in\mathbb{Z}_{>0}}$ *is increasing.*

Proof We prove part (ii) first, with (i) then following easily. Thus suppose that $f(x) \geq 0$ for each $x \in X$. Let $k \in \mathbb{Z}_{>0}$. For $j \in \{1, \dots, k2^k\}$, define
$$A_{k,j} = \{x \in X \mid 2^{-k}(j-1) \leq f(x) < 2^{-k}j\}.$$
As f is measurable, each of these sets is measurable (why?). We then define $f_k(x)$ by
$$f_k(x) = \begin{cases} 2^{-k}(j-1), & x \in A_{k,j}, \\ k, & x \in A \setminus (\cup_{j=1}^{k2^k} A_{k,j}). \end{cases}$$
If $f(x) < \infty$ then the sequence $(f_k(x))_{k\in\mathbb{Z}_{>0}}$ converges to $f(x)$ by construction. If $f(x) = \infty$ then $f_k(x) = k$ for all $k \in \mathbb{Z}_{>0}$, and again the sequence converges, i.e., diverges to ∞.

This proves the result when f is positive-valued. If f is not positive-valued, then one writes $f = f_+ - f_-$ where
$$f_+(x) = \max\{f(x), 0\}, \quad f_-(x) = \max\{-f(x), 0\},$$
cf. Corollary 5.6.17. In this case, the preceding argument can be applied to f_+ and f_- separately, giving (i). ■

One can also consider simple functions that are $\mathbb{C}$- or $\mathbb{R}^n$-valued. Let us first consider the vector-valued case.

5.6.40 Definition (Vector-valued simple function) For a measurable space $(X, \mathscr{A})$, a function $\boldsymbol{f}\colon X \to \mathbb{R}^n$ is a ***simple function*** if each of its components $f_j\colon X \to \mathbb{R}$, $j \in \{1, \dots, n\}$, is a simple function. •

The following characterisation of $\mathbb{R}^n$-valued simple functions is then useful.

5.6.41 Proposition (Characterisation of vector-valued simple functions) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and for* $\mathbf{f}\colon \mathsf{X} \to \mathbb{R}^{\mathsf{n}}$*, the following statements are equivalent:*

(i) $\mathbf{f}$ *is a simple function;*

(ii) $\operatorname{image}(\mathbf{f}) = \{\mathbf{a}_1, \dots, \mathbf{a}_{\mathsf{k}}\} \subseteq \mathbb{R}^{\mathsf{n}}$ *and the sets* $\mathbf{f}^{-1}(\mathbf{a}_{\mathsf{j}})$*,* $\mathsf{j} \in \{1, \dots, \mathsf{k}\}$*, are measurable;*

(iii) there exists $\mathsf{B}_1, \dots, \mathsf{B}_{\mathsf{m}} \in \mathscr{A}$ *and* $\mathbf{b}_1, \dots, \mathbf{b}_{\mathsf{m}} \in \mathbb{R}^{\mathsf{n}}$ *such that* $\mathbf{f} = \sum_{\mathsf{j}=1}^{\mathsf{m}} \mathbf{b}_{\mathsf{j}}\chi_{\mathsf{B}_{\mathsf{j}}}$*;*

(iv) there exist pairwise disjoint sets $\mathsf{C}_1, \dots, \mathsf{C}_{\mathsf{r}} \in \mathscr{A}$ *and* $\mathbf{c}_1, \dots, \mathbf{c}_{\mathsf{r}} \in \mathbb{R}^{\mathsf{n}}$ *such that* $\mathbf{f} = \sum_{\mathsf{j}=1}^{\mathsf{r}} \mathbf{c}_{\mathsf{j}}\chi_{\mathsf{C}_{\mathsf{j}}}$.

Proof It suffices to show the equivalence of any of the last three statements with the first. The arguments from Proposition 5.6.37 can then be applied to show the equivalence with the other two statements, the only difference being the replacement of $\overline{\mathbb{R}}$ with $\mathbb{R}^n$. We shall show that the first statement is equivalent to the fourth.

First suppose that f is a simple function and write

$$f_j = \sum_{k=1}^{r_j} c_{j,k}\chi_{C_{j,k}},$$

for $c_{j,k} \in \mathbb{R}$ and for pairwise disjoint sets $C_{j,k} \in \mathscr{A}$, $j \in \{1, \dots, n\}$, $k \in \{1, \dots, r_j\}$. Let $x \in X$ and denote

$$C(x) = \cap\{C_{j,k} \mid j \in \{1, \dots, n\},\ k \in \{1, \dots, r_j\},\ x \in C_{j,k}\}.$$

Since there are finitely many sets $C_{j,k}$, $j \in \{1, \dots, n\}$, $k \in \{1, \dots, r_j\}$, it follows that there are finitely many possible intersections of these sets. Therefore, there are pairwise disjoint measurable sets $C_1, \dots, C_r$ such that $\{C(x)\}_{x \in X} = \{C_1, \dots, C_r\}$. Moreover, if $x \in X$ and if $f(x) \neq \mathbf{0}$, then $x \in C_l$ for some $l \in \{1, \dots, r\}$. Moreover, since $C_l = C_{j_1,k_1} \cap \dots \cap C_{j_m,k_m}$ for some distinct $j_1, \dots, j_m \in \{1, \dots, n\}$ and some $k_l \in \{1, \dots, r_{j_l}\}$, $l \in \{1 \dots, m\}$, we have

$$f_j(x) = \begin{cases} c_{j_l,k_l}, & j = j_l \text{ for some } l \in \{1, \dots, m\}, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, taking $\boldsymbol{c}_l$ to be the vector whose jth component is given by the expression on the right above, we have

$$f = \sum_{l=1}^{r} \boldsymbol{c}_l\chi_{C_l},$$

as desired.

Conversely, suppose that f satisfies the fourth condition with

$$f = \sum_{l=1}^{r} \boldsymbol{c}_l\chi_{C_l}.$$

Then

$$f_j = \sum_{l=1}^{r} c_{l,j}\chi_{C_l},$$

where $c_{l,j}$, $j \in \{1, \dots, n\}$, is the jth component of $\boldsymbol{c}_l$, $l \in \{1, \dots, r\}$. This shows that f is a simple function. ■

The same constructions obviously apply to $\mathbb{C}$-valued functions, and we record the constructions here.

5.6.42 Definition ($\mathbb{C}$-valued simple function) For a measurable space $(X, \mathscr{A})$, a function $f\colon X \to \mathbb{C}$ is a ***simple function*** if $\mathrm{Re}(f), \mathrm{Im}(f)\colon X \to \mathbb{R}$ are simple functions. •

5.6.43 Corollary (Characterisation of $\mathbb{C}$-valued simple functions) *For a measurable space* $(X, \mathscr{A})$ *and for* $f\colon X \to \mathbb{C}$*, the following statements are equivalent:*

(i) f *is a simple function;*

(ii) $\mathrm{image}(f) = \{a_1, \dots, a_k\} \subseteq \mathbb{C}$ *and the sets* $f^{-1}(a_j)$, $j \in \{1, \dots, k\}$*, are measurable;*

(iii) there exists $B_1, \dots, B_m \in \mathscr{A}$ *and* $b_1, \dots, b_m \in \mathbb{C}$ *such that* $f = \sum_{j=1}^{m} b_j \chi_{B_j}$*;*

(iv) there exist pairwise disjoint sets $C_1, \dots, C_r \in \mathscr{A}$ *and* $c_1, \dots, c_r \in \mathbb{C}$ *such that* $f = \sum_{j=1}^{r} c_j \chi_{C_j}$.

One can also use $\mathbb{C}$- or vector-valued simple functions to approximate $\mathbb{C}$- or vector-valued measurable functions.

5.6.44 Proposition (Approximation of vector-valued measurable functions by simple functions) *If* $(X, \mathscr{A})$ *is a measure space and if* $\mathbf{f}\colon X \to \mathbb{R}^n$ *is measurable, then there exists a sequence* $(\mathbf{f}_k)_{k\in\mathbb{Z}_{>0}}$ *of* $\mathbb{R}^n$*-valued simple functions such that*

(i) $\lim_{k\to\infty} \mathbf{f}_k(x) = \mathbf{f}(x)$ *for each* $x \in X$ *and*

(ii) $\|\mathbf{f}_k(x)\|_{\mathbb{R}^n} \le \|\mathbf{f}(x)\|_{\mathbb{R}^n}$ *for each* $x \in X$.

Proof Let $f_1, \dots, f_n$ be the components of $\boldsymbol{f}$. For each $j \in \{1, \dots, n\}$, if we apply the construction of Proposition 5.6.39, we arrive at a sequence $(f_{j,k})_{k\in\mathbb{Z}_{>0}}$ of simple functions for which

1. $\lim_{k\to\infty} f_{j,k}(x) = f_j(x)$ for every $x \in X$ and
2. $|f_{j,k}(x)| \le |f_j(x)|$ for every $x \in X$

(the verification of the second property requires looking for a moment at the particular construction of Proposition 5.6.39. If we take

$$\boldsymbol{f}_k(x) = (f_{1,k}(x), \dots, f_{n,k}(x)), \qquad x \in X,\ k \in \mathbb{Z}_{>0}$$

then one sees easily that the sequence $(\boldsymbol{f}_k)_{k\in\mathbb{Z}_{>0}}$ has the desired properties. ■

Of course, this specialises to the $\mathbb{C}$ case.

5.6.45 Corollary (Approximation of $\mathbb{C}$-valued measurable functions by simple functions) *If* $(X, \mathscr{A})$ *is a measure space and if* $f\colon X \to \mathbb{C}$ *is measurable, then there exists a sequence* $(f_k)_{k\in\mathbb{Z}_{>0}}$ *of* $\mathbb{C}$*-valued simple functions such that*

(i) $\lim_{k\to\infty} f_k(x) = f(x)$ *for each* $x \in X$ *and*

(ii) $|f_k(x)| \le |f(x)|$ *for each* $x \in X$.

5.6.6 Topological characterisations of convergence for sequences of measurable functions[8]

In this section we characterise some of the modes of convergence for sequences of measurable functions in terms of topological constructions. We let $(X, \mathscr{A}, \mu)$

[8] The results in this section are not used in an essential way elsewhere in the text, except in Sections 5.7.5 and 5.9.11.

be a measure space. It will be useful to characterise measurable functions as equivalence classes of functions that agree up to sets of measure zero. Thus we say that $f, g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$ are ***equivalent*** if

$$\mu(\{x \in X \mid f(x) \neq g(x)\}) = 0.$$

This is readily seen to define an equivalence relation in $\mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$ and we denote by $\mathsf{L}^{0}((X, \mathscr{A}); \overline{\mathbb{R}})$ the set of equivalence classes, an equivalence class being denoted by $[f]$ for $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$. The following result shows that convergence pointwise almost everywhere is defined independently of equivalence classes.

5.6.46 Lemma (Almost everywhere pointwise convergence is independent of equivalence) *Let* $(X, \mathscr{A}, \mu)$ *be a measure space. For a sequence* $([f_j])_{j \in \mathbb{Z}_{>0}}$ *in* $\mathsf{L}^{0}((X, \mathscr{A}); \mathbb{R})$ *and for* $[f] \in \mathsf{L}^{0}((X, \mathscr{A}); \mathbb{R})$ *the following statements are equivalent:*

(i) *there exists a sequence* $(g_j)_{j \in \mathbb{Z}_{>0}}$ *in* $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ *and* $g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ *such that*

  (a) $[g_j] = [f_j]$ *for* $j \in \mathbb{Z}_{>0}$,

  (b) $[g] = [f]$, *and*

  (c) $(g_j)_{j \in \mathbb{Z}_{>0}}$ *converges pointwise almost everywhere to* g.

(ii) *for every sequence* $(g_j)_{j \in \mathbb{Z}_{>0}}$ *in* $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ *and for every* $g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ *satisfying*

  (a) $[g_j] = [f_j]$ *for* $j \in \mathbb{Z}_{>0}$ *and*

  (b) $[g] = [f]$,

  it holds that $(g_j)_{j \in \mathbb{Z}_{>0}}$ *converges pointwise almost everywhere to* g.

Proof It is clear that the second statement implies the first, so we only prove the converse. Thus we let $(g_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ and $g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ be such that

1. $[g_j] = [f_j]$ for $j \in \mathbb{Z}_{>0}$,
2. $[g] = [f]$, and
3. $(g_j)_{j \in \mathbb{Z}_{>0}}$ converges pointwise almost everywhere to g.

Let $(h_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ and let $h \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ be such that

1. $[h_j] = [f_j]$ for $j \in \mathbb{Z}_{>0}$ and
2. $[h] = [f]$.

Define

$$A = \{x \in X \mid g(x) \neq f(x)\}, \quad B = \{x \in X \mid h(x) \neq f(x)\}$$

and, for $j \in \mathbb{Z}_{>0}$, define

$$A_j = \{x \in X \mid g_j(x) \neq f_j(x)\}, \quad B_j = \{x \in X \mid h_j(x) \neq f_j(x)\}$$

and note that

$$x \in X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B) \quad \Longrightarrow \quad h(x) = f(x) = g(x)$$

and

$$x \in X \setminus (A_j \cup B_j) = (X \setminus A_j) \cap (X \setminus B_j) \quad \Longrightarrow \quad h_j(x) = f_j(x) = g_j(x).$$

Thus,

$$x \in X \setminus \Big((\cup_{j \in \mathbb{Z}_{>0}} A_j \cup B_j) \cup (A \cup B)\Big) \quad \Longrightarrow \quad \lim_{j\to\infty} h_j(x) = \lim_{j\to\infty} g_j(x) = g(x) = h(x).$$

Since $(\cup_{j \in \mathbb{Z}_{>0}} A_j \cup B_j) \cup (A \cup B)$ is a countable union of sets of measure zero, it has zero measure, and so $(h_j)_{j \in \mathbb{Z}_{>0}}$ converges pointwise almost everywhere to h. ■

With the preceding lemma, the following definition makes sense.

5.6.47 Definition (Almost everywhere convergence of sequences of equivalence classes of functions) Let $(X, \mathscr{A}, \mu)$ be a measure space, let $([f_j])_{j \in \mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^0((X, \mathscr{A}); \overline{\mathbb{R}})$, and let $[f] \in \mathsf{L}^0((X, \mathscr{A}); \overline{\mathbb{R}})$. The sequence $([f_j])_{j \in \mathbb{Z}_{>0}}$ ***converges pointwise almost everywhere*** to $[f]$ if

$$\mu\Big(X \setminus \Big\{x \in X \;\Big|\; f(x) = \lim_{j\to\infty} f_j(x)\Big\}\Big) = 0. \qquad \bullet$$

We begin by indicating that the convergence defined by almost everywhere pointwise convergence cannot arise from a topology.

5.6.48 Proposition (Almost everywhere pointwise convergence is not always topological) *Let* $(\mathrm{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathscr{T}_{\mathrm{a.e.}}$ *be the set of topologies* τ *on* $\mathsf{L}^0((\mathrm{X}, \mathscr{A}); \mathbb{R})$ *such that the convergent sequences in* τ *are precisely the almost everywhere pointwise convergent sequences. If there exists a sequence* $(\mathrm{f_j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *in* $\mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \mathbb{R})$ *and* $\mathrm{f} \in \mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \mathbb{R})$ *such that* $(\mathrm{f_j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *converges in measure to* f *but does not converge pointwise almost everywhere to* f*, then* $\mathscr{T}_{\mathrm{a.e.}} = \emptyset$*.*

Proof Let us denote by $z \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ the zero function. The hypotheses ensure that the sequence $(g_j \triangleq f_j - f)_{j \in \mathbb{Z}_{>0}}$ converges to z in measure, but does not converges pointwise almost everywhere to z. Suppose that $\mathscr{T}_{\mathrm{a.e.}} \neq \emptyset$ and let $\tau \in \mathscr{T}_{\mathrm{a.e.}}$. Since almost everywhere pointwise convergence agrees with convergence in τ, there exists a neighbourhood U of $[z]$ in $\mathsf{L}^0((X, \mathscr{A}); \mathbb{R})$ such that the set

$$\{j \in \mathbb{Z}_{>0} \mid [f_j] \in U\}$$

is finite. By Proposition 5.6.26 there exists a subsequence $(f_{j_k})_{k \in \mathbb{Z}_{>0}}$ of $(f_j)_{j \in \mathbb{Z}_{>0}}$ that converges pointwise almost everywhere to z. Thus the sequence $([f_{j_k}])_{k \in \mathbb{Z}_{>0}}$ converges pointwise almost everywhere to $[z]$, and so converges to $[z]$ in τ. Thus, in particular, the set

$$\{k \in \mathbb{Z}_{>0} \mid [f_{j_k}] \in U\}$$

is infinite, which is a contradiction. ■

In particular, we have the following result which shows that in the most common situation where one wishes to study almost everywhere pointwise convergence, this sort of convergence is not topological.

5.6.49 Corollary (Almost everywhere pointwise convergence is not topological for the Lebesgue measure) *Let $\mathscr{T}_{\mathrm{a.e.}}$ be the set of topologies τ on $\mathsf{L}^0((\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n)); \mathbb{R})$ such that the convergent sequences in τ are precisely the almost everywhere pointwise convergent sequences using the Lebesgue measure on $\mathbb{R}^n$. Then $\mathscr{T}_{\mathrm{a.e.}} = \emptyset$.*

Proof In Example 5.6.25 we have seen that there exists a sequence in $\mathsf{L}^{(0)}((\mathbb{R}, \mathscr{L}(\mathbb{R})); \mathbb{R})$ that converges in measure but does not converge pointwise almost everywhere. This example is easily adapted to $\mathsf{L}^{(0)}((\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n)); \mathbb{R})$, and the result then follows from Proposition 5.6.48. ■

Now one can ask if there is a framework in which almost everywhere pointwise convergence can be studied. Indeed there is such a framework. The construction relies on notions concerning filters and nets from ***missing stuff***.

5.6.50 Definition (Limit structure) A ***limit structure*** on a set S is a subset $\mathscr{L} \subseteq \mathscr{F}(S) \times S$ with the following properties:

(i) if $x \in S$ then $(\mathcal{F}_x, x) \in \mathscr{L}$;

(ii) if $(\mathcal{F}, x) \in \mathscr{L}$ and if $\mathcal{F} \subseteq \mathcal{G} \in \mathscr{F}(S)$ then $(\mathcal{G}, x) \in \mathscr{L}$;

(iii) if $(\mathcal{F}, x), (\mathcal{G}, x) \in \mathscr{L}$ then $(\mathcal{F} \cap \mathcal{G}, x) \in \mathscr{L}$.

If $(\Lambda, \preceq)$ is a directed set, a Λ-net $\phi \colon \Lambda \to S$ is ***$\mathscr{L}$-convergent*** to $x \in S$ if $(\mathcal{F}_\phi, x) \in \mathscr{L}$. Let us denote by $\mathscr{S}(\mathscr{L})$ the set of $\mathscr{L}$-convergent $\mathbb{Z}_{>0}$-nets, i.e., the set of $\mathscr{L}$-convergent sequences. •

The intuition behind the notion of a limit structure is as follows. Condition (i) says that the trivial filter converging to x should be included in the limit structure, condition (ii) says that if a filter converges to x, then every coarser filter also converges to x, and condition (iii) says that "mixing" filters converging to x should give a filter converging to x. Starting from the definition of a limit structure, one can reproduce many of the concepts from topology, e.g., openness, closedness, compactness, continuity.

We are interested in the special case of limit structures on a vector space V. We suppose that V is defined over a field F. For $\mathcal{F}, \mathcal{G} \in \mathscr{F}(\mathsf{V})$ and for $a \in \mathsf{F}$ we denote

$$\mathcal{F} + \mathcal{G} = \{A + B \mid A \in \mathcal{F},\ B \in \mathcal{G}\}, \quad a\mathcal{F} = \{aA \mid A \in \mathcal{F}\},$$

where, as usual,

$$A + B = \{u + v \mid u \in A,\ v \in B\}, \quad aA = \{au \mid u \in A\}.$$

We say that a limit structure $\mathscr{L}$ on a vector space V is ***linear*** if $(\mathcal{F}_1, v_1), (\mathcal{F}_2, v_2) \in \mathscr{L}$ implies that $(\mathcal{F}_1 + \mathcal{F}_2, v_1 + v_2) \in \mathscr{L}$ and if $a \in \mathsf{F}$ and $(\mathcal{F}, v) \in \mathscr{L}$ then $(a\mathcal{F}, av) \in \mathscr{L}$.

For $[f] \in \mathsf{L}^0((X, \mathscr{A}); \mathbb{R})$ define

$$\mathscr{F}_{[f]} = \{\mathcal{F} \in \mathscr{F}(\mathsf{L}^0((X, \mathscr{A}); \mathbb{R})) \mid \mathcal{F}_\phi \subseteq \mathcal{F} \text{ for some } \mathbb{Z}_{>0}\text{-net } \phi \text{ such that } (\phi(j))_{j \in \mathbb{Z}_{>0}} \text{ is almost everywhere pointwise convergent to } [f]\}.$$

We may now define a limit structure on $\mathsf{L}^0((X, \mathscr{A}); \mathbb{R})$ as follows.

5.6.51 Theorem (Almost everywhere pointwise convergence is defined by a limit structure) *The subset of* $\mathscr{F}(\mathsf{L}^0((\mathsf{X},\mathscr{A});\mathbb{R}))\times\mathsf{L}^0((\mathsf{X},\mathscr{A});\mathbb{R})$ *defined by*

$$\mathscr{L}_\mu = \{(\mathcal{F},[\mathrm{f}]) \mid \mathcal{F}\in\mathscr{F}_{[\mathrm{f}]}\}$$

is a linear limit structure on $\mathsf{L}^0((\mathsf{X},\mathscr{A});\mathbb{R})$*. Moreover, a sequence* $([\mathrm{f_j}])_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is* $\mathscr{L}_\mu$*-convergent to* $[\mathrm{f}]$ *if and only if the sequence is almost everywhere pointwise convergent to* $[\mathrm{f}]$*.*

Proof Let $[f]\in\mathsf{L}^0((X,\mathscr{A});\mathbb{R})$. Consider the trivial $\mathbb{Z}_{>0}$-net $\phi_{[f]}\colon\mathbb{Z}_{>0}\to\mathsf{L}^0((X,\mathscr{A});\mathbb{R})$ defined by $\phi_{[f]}(j)=[f]$. Since $\mathcal{F}_\phi=\mathcal{F}_{[f]}$ and since $(\mathcal{F}_\phi,[f])\in\mathscr{L}_\mu$, the condition (i) for a limit structure is satisfied.

Let $(\mathcal{F},[f])\in\mathscr{L}_\mu$ and suppose that $\mathcal{F}\subseteq\mathcal{G}$. Then $\mathcal{F}\in\mathscr{F}_{[f]}$ and so $\mathcal{F}\supseteq\mathcal{F}_\phi$ for some $\mathbb{Z}_{>0}$-net ϕ that converges pointwise almost everywhere to $[f]$. Therefore, we immediately have $\mathcal{F}_\phi\subseteq\mathcal{G}$ and so $(\mathcal{G},[f])\in\mathscr{L}_\mu$. This verifies condition (ii) in the definition of a limit structure.

Finally, let $(\mathcal{F},[f]),(\mathcal{G},[f])\in\mathscr{L}_\mu$ and let ϕ and ψ be $\mathbb{Z}_{>0}$-nets that converge pointwise almost everywhere to $[f]$ and satisfy $\mathcal{F}_\phi\subseteq\mathcal{F}$ and $\mathcal{F}_\psi\subseteq\mathcal{G}$. Define a $\mathbb{Z}_{>0}$-net $\phi\wedge\psi$ by

$$\phi\wedge\psi(j)=\begin{cases}\phi(\frac{1}{2}(j+1)), & j\text{ odd},\\ \psi(\frac{1}{2}j), & j\text{ even}.\end{cases}$$

We first claim that $\phi\wedge\psi$ converges pointwise almost everywhere to $[f]$. Let

$$A=\Big\{x\in X\ \Big|\ \lim_{j\to\infty}\phi(j)(x)\neq f(x)\Big\},\quad B=\Big\{x\in X\ \Big|\ \lim_{j\to\infty}\psi(j)(x)\neq f(x)\Big\}.$$

If $x\in X\setminus(A\cup B)$ then

$$\lim_{j\to\infty}\phi(j)(x)=\lim_{j\to\infty}\psi(j)(x)=f(x).$$

Thus, for $x\in X\setminus(A\cup B)$ and $\epsilon\in\mathbb{R}_{>0}$ there exists $N\in\mathbb{Z}_{>0}$ such that

$$|f(x)-\phi(j)(x)|,|f(x)-\psi(j)(x)|<\epsilon,\qquad j\geq N.$$

Therefore, for $j\geq 2N$ and for $x\in X\setminus(A\cup B)$ we have $|f(x)-\phi\wedge\psi(j)(x)|<\epsilon$ and so

$$\lim_{j\to\infty}\phi\wedge\psi(j)(x)=f(x),\qquad x\in X\setminus(A\cup B).$$

Since $\mu(A\cup B)=0$ it indeed follows that $\phi\wedge\psi$ converges pointwise almost everywhere to $[f]$.

We next claim that $\mathcal{F}_{\phi\wedge\psi}\subseteq\mathcal{F}\cap\mathcal{G}$. Indeed, let $S\in\mathcal{F}_{\phi\wedge\psi}$. Then there exists $N\in\mathbb{Z}_{>0}$ such that $T_{\phi\wedge\psi}(N)\subseteq S$. Therefore, there exists $N_\phi,N_\psi\in\mathbb{Z}_{>0}$ such that $T_\phi(N_\phi)\subseteq S$ and $T_\psi(N_\psi)\subseteq S$. That is, $S\in\mathcal{F}_\phi\cap\mathcal{F}_\psi\subseteq\mathcal{F}\cap\mathcal{G}$. This shows that $(\mathcal{F}\cap\mathcal{G},[f])\in\mathscr{L}_\mu$ and so shows that condition (iii) in the definition of a limit structure holds.

Thus we have shown that $\mathscr{L}_\mu$ is a limit structure. Let us show that it is a linear limit structure. Let $(\mathcal{F}_1,[f_1]),(\mathcal{F}_2,v_2)\in\mathscr{L}_\mu$. Thus there exists $\mathbb{Z}$-nets ϕ_1 and ϕ_2 in $\mathsf{L}^0((X,\mathscr{A});\mathbb{R})$ converging pointwise almost everywhere to $[f_1]$ and $[f_2]$, respectively, and such that $\mathcal{F}_{\phi_1}\subseteq\mathcal{F}_1$ and $\mathcal{F}_{\phi_2}\subseteq\mathcal{F}_2$. Let us denote by $(f_{1,j})_{j\in\mathbb{Z}_{>0}}$ and $(f_{2,j})_{j\in\mathbb{Z}_{>0}}$ sequences in $\mathsf{L}^{(0)}((X,\mathscr{A});\mathbb{R})$ such that $[f_{1,j}]=\phi_1(j)$ and $[f_{2,j}]=\phi_2(j)$ for $j\in\mathbb{Z}_{>0}$. Then, as in the proof of Lemma 5.6.46, there exists a subset $A\subseteq X$ of zero measure such that

$$\lim_{j\to\infty}f_{j,1}(x)=f_1(x),\quad \lim_{j\to\infty}f_{2,j}(x)=f_2(x),\qquad x\in X\setminus A.$$

Thus, for $x \in X \setminus A$,

$$\lim_{j\to\infty}(f_{1,j} + f_{2,j})(x) = (f_1 + f_2)(x).$$

This shows that the $\mathbb{Z}_{>0}$-net $\phi_1 + \phi_2$ converges pointwise almost everywhere to $[f_1 + f_2]$. Since $\mathcal{F}_{\phi_1+\phi_2} \subseteq \mathcal{F}_1 + \mathcal{F}_2$, it follows that $(\mathcal{F}_1 + \mathcal{F}_2, [f_1 + f_2]) \in \mathscr{L}_\mu$. An entirely similarly styled argument gives $(a\mathcal{F}, av) \in \mathscr{L}_\mu$ for $(\mathcal{F}, v) \in \mathscr{L}_\mu$.

We now need to show that $\mathscr{S}(\mathscr{L}_\mu)$ consists exactly of the almost everywhere pointwise convergent sequences. The very definition of $\mathscr{L}_\mu$ ensures that if a $\mathbb{Z}_{>0}$-net ϕ is almost everywhere pointwise convergent then $\phi \in \mathscr{S}(\mathscr{L}_\mu)$. We prove the converse, and so let ϕ be $\mathscr{L}_\mu$-convergent to $[f]$. Therefore, by definition of $\mathscr{L}_\mu$, there exists a $\mathbb{Z}_{>0}$-net ψ converging pointwise almost everywhere to $[f]$ such that $\mathcal{F}_\psi \subseteq \mathcal{F}_\phi$.

1 **Lemma** *There exists of a subsequence ψ' of ψ such that $\mathcal{F}_{\psi'} = \mathcal{F}_\phi$.*

Proof Let $n \in \mathbb{Z}_{>0}$ and note that $T_\psi(n) \in \mathcal{F}_\psi \subseteq \mathcal{F}_\phi$. Thus there exists $k \in \mathbb{Z}_{>0}$ such that $T_\phi(k) \subseteq T_\psi(n)$. Then define

$$k_n = \min\{k \in \mathbb{Z}_{>0} \mid T_\phi(k) \subseteq T_\psi(n)\},$$

the minimum being well-defined since

$$k > k' \quad\implies\quad T_\phi(k) \subseteq T_\phi(k').$$

This uniquely defines, therefore, a sequence $(k_n)_{n\in\mathbb{Z}_{>0}}$. Moreover, if $n_1 > n_2$ then $T_\psi(n_2) \subseteq T_\psi(n_1)$ which implies that $T_\phi(k_{n_2}) \subseteq T_\psi(n_1)$. Therefore, $k_{n_2} \geq k_{n_1}$, showing that the sequence $(k_n)_{n\in\mathbb{Z}_{>0}}$ is nondecreasing.

Now define $\theta\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ as follows. If $j < k_n$ for every $n \in \mathbb{Z}_{>0}$ then define $\theta(j)$ in an arbitrary manner. If $j \geq k_1$ then note that $\phi(j) \in T_\phi(k_1) \subseteq T_\psi(1)$. Thus there exists (possibly many) $m \in \mathbb{Z}_{>0}$ such that $\phi(j) = \psi(m)$. If $j \geq k_n$ for $n \in \mathbb{Z}_{>0}$ then there exists (possibly many) $m \geq n$ such that $\phi(j) = \psi(m)$. Thus for any $j \in \mathbb{Z}_{>0}$ we can define $\theta(j) \in \mathbb{Z}_{>0}$ such that $\phi(j) = \psi(\theta(j))$ if $j \geq k_1$ and such that $\theta(j) \geq n$ if $j \geq k_n$.

Note that any function $\theta\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ as constructed above is unbounded. Therefore, there exists a strictly increasing function $\rho\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ such that $\operatorname{image}(\rho) = \operatorname{image}(\theta)$. We claim that $\mathcal{F}_\rho = \mathcal{F}_\theta$. First let $n \in \mathbb{Z}_{>0}$ and let $j \geq k_{\rho(n)}$. Then $\theta(j) \geq \rho(n)$. Since $\operatorname{image}(\rho) = \operatorname{image}(\theta)$ there exists $m \in \mathbb{Z}_{>0}$ such that $\rho(m) = \theta(j) \geq \rho(n)$. Since ρ is strictly increasing, $m \geq n$. Thus $\theta(j) \in T_\rho(n)$ and so $T_\theta(k_{\rho(n)}) \subseteq T_\rho(n)$. This implies that $\mathcal{F}_\rho \subseteq \mathcal{F}_\theta$.

Conversely, let $n \in \mathbb{Z}_{>0}$ and let $r_n \in \mathbb{Z}_{>0}$ be such that

$$\rho(r_n) > \max\{\theta(1), \dots, \theta(n)\};$$

this is possible since ρ is unbounded. If $j \geq r_n$ then

$$\rho(j) \geq \rho(r_n) > \max\{\theta(1), \dots, \theta(n)\}.$$

Since $\operatorname{image}(\rho) = \operatorname{image}(\theta)$ we have $\rho(j) = \theta(m)$ for some $m \in \mathbb{Z}_{>0}$. We must have $m > n$ and so $\rho(j) \in T_\theta(n)$. Thus $T_\rho(r_n) \subseteq T_\theta(n)$ and so $\mathcal{F}_\theta \subseteq \mathcal{F}_\rho$.

To arrive at the conclusions of the lemma we first note that, by definition of θ, $\mathcal{F}_\phi = \mathcal{F}_{\psi\circ\theta}$. We now define $\psi' = \psi \circ \rho$ and note that

$$\mathcal{F}_\phi = \mathcal{F}_{\psi\circ\theta} = \psi(\mathcal{F}_\theta) = \psi(\mathcal{F}_\rho) = \mathcal{F}_{\psi\circ\rho},$$

as desired. ▼

Since a subsequence of an almost everywhere pointwise convergent sequence is almost everywhere pointwise convergent to the same limit, it follows that ψ', and so ϕ, converges almost everywhere pointwise to $[f]$. ■

Note that we have already seen in Sections **??** and **??** that pointwise and uniform convergence is prescribed by a topology. We shall see in ***missing stuff*** that convergence in measure is topological.

Exercises

5.6.1 Let $(X, \mathscr{A}, \mu)$ be a measure space that is not complete. Show that Proposition 5.6.10 fails in this case.

5.6.2 Let $(X, \mathscr{A}, \mu)$ be a measure space that is not complete. Show that Corollary 5.6.19 fails in this case.

5.6.3 Let $(X, \mathscr{A})$ be a measurable space and let $A, B \in \mathscr{A}$ be such that $X = A \overset{\circ}{\cup} B$. Let $f_A \colon A \to \overline{\mathbb{R}}$ be $\mathscr{A}_A$-measurable and let $f_B \colon B \to \overline{\mathbb{R}}$ be $\mathscr{A}_B$-measurable. Show that $f \colon X \to \overline{\mathbb{R}}$ defined by

$$f(x) = \begin{cases} f_A(x), & x \in A, \\ f_B(x), & x \in B \end{cases}$$

is $\mathscr{A}$-measurable.

5.6.4 Give an example of a measure space $(X, \mathscr{A}, \mu)$, a sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$, and a function $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$ such that $(f_j)_{j \in \mathbb{Z}_{>0}}$ converges pointwise almost everywhere to f, but does not converges pointwise to f.

Section 5.7

Integration on measure spaces

Up to now, we have studied measurable and measure spaces in some detail. These subjects certainly have some value in their own right, particularly in the domain of probability theory which we discuss in ***missing stuff***. In particular, the properties of the Lebesgue measure on $\mathbb{R}$ and $\mathbb{R}^n$ considered in Sections 5.4 and 5.5 are substantially useful. Following our discussion of measure, we introduced a particular class of functions on measurable spaces called measurable functions. While we showed in Sections 5.9.1 and **??** that for the Lebesgue measure that these functions are not too far from easily understood functions such as step or continuous functions, the importance of measurable functions is perhaps not so easily understood. What we see in this section is that these functions form the basis for a powerful and general theory of integration. For the Lebesgue measure, this construction of the integral generalises the Riemann integral, and repairs some of the defects of the latter as seen in Section 5.1.

The treatment of the integral is as easily carried out in the general setting of a general measure space as it is for the specific case of the Lebesgue integral in particular. Thus we do much of the work in this general setting. In Sections 5.9 and **??** we consider the Lebesgue integral, but only its particular properties that rely on the structure of Lebesgue measure. Thus a reader wanting only to learn about the Lebesgue integral will have to learn it here. A reader only believing they are interested in Lebesgue integration will have to be satisfied by mentally making the replacement of "$(X, \mathscr{A}, \mu)$" with "$(\mathbb{R}, \mathscr{L}(\mathbb{R}), \lambda)$" or "$(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n), \lambda_n)$."

Do I need to read this section? Clearly if you are reading this chapter, then you must read this section. •

5.7.1 Definition of the integral

We consider a measure space $(X, \mathscr{A}, \mu)$. The objective is to define the integral of a measurable function $f\colon X \to \overline{\mathbb{R}}$. We do this in three stages.

Integration of nonnegative simple functions

Let $f \in \mathsf{S}(X; \overline{\mathbb{R}}_{\geq 0})$ be written as $f = \sum_{j=1}^{k} a_j \chi_{A_j}$ for a partition $(A_1, \dots, A_k)$ of X into measurable sets. Let us first make an observation concerning the fact that the numbers $a_1, \dots, a_k$ and the sets $A_1, \dots, A_k$ are not uniquely prescribed by f.

5.7.1 Proposition (Independence of integral of simple functions on partition) *For a measure space $(X, \mathscr{A}, \mu)$ suppose that $f \in \mathsf{S}(X; \overline{\mathbb{R}}_{\geq 0})$ satisfies*

$$f = \sum_{j=1}^{k} a_j \chi_{A_j} = \sum_{l=1}^{m} b_l \chi_{B_l}$$

for $\mathrm{a}_1,\dots,\mathrm{a}_\mathrm{k},\mathrm{b}_1,\dots,\mathrm{b}_\mathrm{m}\in\overline{\mathbb{R}}_{\geq 0}$ *and* $\mathrm{A}_1,\dots,\mathrm{A}_\mathrm{k}\in\mathscr{A}$ *disjoint and* $\mathrm{B}_1,\dots,\mathrm{B}_\mathrm{m}\in\mathscr{A}$ *disjoint. Then*

$$\sum_{\mathrm{j}=1}^{\mathrm{k}}\mathrm{a}_\mathrm{j}\mu(\mathrm{A}_\mathrm{j})=\sum_{\mathrm{l}=1}^{\mathrm{m}}\mathrm{b}_\mathrm{l}\mu(\mathrm{B}_\mathrm{l}).$$

Proof Without loss of generality we suppose that none of $a_1,\dots,a_k$ and $b_1,\dots,b_m$ are zero. It therefore follows that $\cup_{j=1}^k A_j=\cup_{l=1}^m B_l$. Note that if $A_j\cap B_m\neq\emptyset$ for some $j\in\{1,\dots,k\}$ and $l\in\{1,\dots,m\}$, it follows that $a_j=b_l$. Therefore, we have

$$\sum_{j=1}^k a_j\mu(A_j)=\sum_{j=1}^k\sum_{l=1}^m a_j\mu(A_j\cap B_l)=\sum_{l=1}^m\sum_{j=1}^k b_l\mu(B_l\cap A_j)=\sum_{l=1}^m b_l\mu(B_l),$$

as desired. ■

Given the preceding result, the following definition makes sense.

5.7.2 Definition (Integral of nonnegative simple function) For a measure space $(X,\mathscr{A},\mu)$ and for $f\in\mathsf{S}(X;\overline{\mathbb{R}}_{\geq 0})$ given by $f=\sum_{j=1}^k a_j\chi_{A_j}$ for a partition $(A_1,\dots,A_k)$ of X into measurable sets, the ***integral*** of f is

$$\int_X f\,\mathrm{d}\mu=\sum_{j=1}^k a_j\mu(A_j).$$ •

Note that the notion of integral for a simple function is a natural adaptation of the notion of integral for a step function in our development of the Riemann integral in Sections 3.4 and **??**.

Let us give some examples of simple functions and their integrals.

5.7.3 Examples (Positive simple functions and their integrals)

1. Let $P=(I_1,\dots,I_k)$ be a partition of $[a,b]\subseteq\mathbb{R}$ with endpoints $\mathrm{EP}(P)=(x_0,x_1,\dots,x_k)$ and let $f\colon[a,b]\to\mathbb{R}$ be a step function taking value c_j on the interval I_j, $j\in\{1,\dots,k\}$. Clearly then, f is also a simple function since intervals are measurable. Moreover,
$$\int_{[a,b]}f\,\mathrm{d}\lambda=\int_a^b f(x)\,\mathrm{d}x=\sum_{j=1}^k c_j(x_j-x_{j-1}),$$
since the Lebesgue measure of an interval is its length.
2. Let us consider the measure space $(\mathbb{R},\mathscr{L}(\mathbb{R}),\lambda)$ and take $A=\mathbb{Q}$. By Exercise 2.5.8 it follows that $\lambda(A)=0$. Therefore, the simple function χ_A has measure zero.
3. Let X be a set, let $\mathscr{A}=\mathbf{2}^X$, and let μ_Σ be the counting measure on X; see Example 5.3.9–3. Let $A_1,\dots,A_k\subseteq X$ be nonempty disjoint subsets, let $a_1,\dots,a_k\in\overline{\mathbb{R}}_{\geq 0}$, and define $f=\sum_{j=1}^k a_j\chi_{A_j}$. If $\mathrm{card}(A_j)=\infty$ for any $j\in\{1,\dots,k\}$ for which $a_j\neq 0$ or if $a_j=\infty$ for any $j\in\{1,\dots,k\}$, then $\int_X f\,\mathrm{d}\mu_\Sigma=\infty$. Otherwise,
$$\int_X f\,\mathrm{d}\mu_\Sigma=\sum_{j=1}^k a_j\,\mathrm{card}(A_j).$$ •

Integration of nonnegative measurable functions

Using the definition of the integral for simple functions, it is possible to immediately deduce a definition of the integral for nonnegative-valued functions. This is done as follows.

5.7.4 Definition (Integral of a nonnegative measurable function) For a measure space $(X, \mathscr{A}, \mu)$ and for $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$, the ***integral*** of f is

$$\int_X f\,\mathrm{d}\mu = \sup\left\{ \int_X g\,\mathrm{d}\mu \;\middle|\; g \in \mathsf{S}(X; \overline{\mathbb{R}}_{\geq 0}) \text{ satisfies } 0 \leq g(x) \leq f(x) \text{ for } x \in X \right\}. \quad \bullet$$

The following result gives a useful characterisation of the integral of nonnegative-valued functions. It also gives an idea of why measurable functions are the "right" class of functions to integrate, since they are well-approximated by simple functions.

5.7.5 Proposition (Sequential characterisation of the integral for nonnegative functions) *Let* $(\mathrm{X}, \mathscr{A}, \mu)$ *be a measure space, let* $\mathrm{f} \in \mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$*, and let* $(\mathrm{f}_\mathrm{j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *be a sequence of increasing positive simple functions converging to* f *as in Proposition 5.6.39. Then*

$$\int_\mathrm{X} \mathrm{f}\,\mathrm{d}\mu = \lim_{\mathrm{j} \to \infty} \int_\mathrm{X} \mathrm{f}_\mathrm{j}\,\mathrm{d}\mu.$$

Proof First we prove the result in the case that f is a simple function.

1 Lemma *Let* $(\mathrm{X}, \mathscr{A}, \mu)$ *be a measure space, let* $\mathrm{f} \in \mathsf{S}(\mathrm{X}; \overline{\mathbb{R}}_{\geq 0})$*, and let* $(\mathrm{f}_\mathrm{j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *be a sequence of increasing positive simple functions converging to* f *as in Proposition 5.6.39. Then*

$$\int_\mathrm{X} \mathrm{f}\,\mathrm{d}\mu = \lim_{\mathrm{j} \to \infty} \int_\mathrm{X} \mathrm{f}_\mathrm{j}\,\mathrm{d}\mu.$$

Proof By Exercise 5.7.1 the sequence $(\int_X f_j\,\mathrm{d}\mu)_{j \in \mathbb{Z}_{>0}}$ is increasing and bounded above by $\int_X f\,\mathrm{d}\mu$. Thus the sequence $(\int_X f_j\,\mathrm{d}\mu)_{j \in \mathbb{Z}_{>0}}$ converges in $\overline{\mathbb{R}}_{\geq 0}$, by Theorem 2.3.8 if the limit is finite, tautologically otherwise. Thus we have

$$\lim_{j \to \infty} \int_X f_j\,\mathrm{d}\mu \leq \int_X f\,\mathrm{d}\mu.$$

Next let $\epsilon \in (0, 1)$. Let us write $f = \sum_{l=1}^m a_l \chi_{A_l}$ for $a_1, \dots, a_m \in \overline{\mathbb{R}}_{\geq 0}$ and disjoint $A_1, \dots, A_m \in \mathscr{A}$. For $l \in \{1, \dots, m\}$ and $j \in \mathbb{Z}_{>0}$ denote

$$A_{j,l} = \{x \in A_l \mid f_j(x) \geq (1 - \epsilon) a_l\},$$

noting that $A_{j,l} \in \mathscr{A}$ since f_j is measurable. Since the sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ is monotonically increasing, the sequence $(A_{j,l})_{j \in \mathbb{Z}_{>0}}$ satisfies

$$A_{j,l} \subseteq A_{j+1,l}, \qquad \cup_{j \in \mathbb{Z}_{>0}} A_{j,l} = A_l.$$

Let us define simple functions

$$g_j = \sum_{l=1}^{m} (1-\epsilon) a_l \chi_{A_{j,l}}, \qquad j \in \mathbb{Z}_{>0}.$$

By Proposition 5.3.3 we have

$$\lim_{j\to\infty} \int_X g_j \,\mathrm{d}\mu = \lim_{j\to\infty} \sum_{l=1}^{m} (1-\epsilon) a_l \mu(A_{j,l}) = \sum_{l=1}^{m} (1-\epsilon) a_l \mu(A_l) = (1-\epsilon) \int_X f \,\mathrm{d}\mu.$$

Since $g_j(x) \le f_j(x)$ for every $j \in \mathbb{Z}_{>0}$, by Exercise 5.7.1 we have

$$\begin{aligned} & \int_X g_j \,\mathrm{d}\mu \le \int_X f_j \,\mathrm{d}\mu \\ \implies \quad & \lim_{j\to\infty} \int_X g_j \,\mathrm{d}\mu \le \lim_{j\to\infty} \int_X f_j \,\mathrm{d}\mu \\ \implies \quad & (1-\epsilon) \int_X f \,\mathrm{d}\mu \le \lim_{j\to\infty} \int_X f_j \,\mathrm{d}\mu \le \int_X f \,\mathrm{d}\mu. \end{aligned}$$

Since ϵ is arbitrary, this implies that

$$\lim_{j\to\infty} \int_X f_j \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu,$$

as desired. ▼

In the case that f is a general nonnegative-valued measurable function, we note that

$$\int_X f_j \,\mathrm{d}\mu \le \int_X f_{j+1} \,\mathrm{d}\mu, \qquad j \in \mathbb{Z}_{>0},$$

and

$$\int_X f_j \,\mathrm{d}\mu \le \int_X f \,\mathrm{d}\mu, \qquad j \in \mathbb{Z}_{>0}.$$

Thus the sequence $(\int_X f_j \,\mathrm{d}\mu)_{j\in\mathbb{Z}_{>0}}$ converges in $\overline{\mathbb{R}}_{\ge 0}$ to a limit (by Theorem 2.3.8 if the limit is finite, tautologically otherwise) and this limit satisfies

$$\lim_{j\to\infty} \int_X f_j \,\mathrm{d}\mu \le \int_X f \,\mathrm{d}\mu.$$

Next let $\epsilon \in \mathbb{R}_{>0}$ and let $g \in \mathsf{S}(X; \overline{\mathbb{R}}_{\ge 0})$ be such that

$$\int_X g \,\mathrm{d}\mu \ge (1-\epsilon) \int_X f \,\mathrm{d}\mu.$$

Define $g_j(x) = \min\{g(x), f_j(x)\}$, and note that g_j is a nonnegative simple function, and that the sequence $(g_j(x))_{j\in\mathbb{Z}_{>0}}$ converges to $g(x)$ for each $x \in X$. By the lemma above we thus have

$$\lim_{j\to\infty} \int_X g_j \,\mathrm{d}\mu = \int_X g \,\mathrm{d}\mu.$$

By Exercise 5.7.1 we have

$$\int_X g_j \, \mathrm{d}\mu \leq \int_X f_j \, \mathrm{d}\mu \quad \Longrightarrow \quad \int_X g \, \mathrm{d}\mu \leq \lim_{j \to \infty} \int_X f_j \, \mathrm{d}\mu$$

which gives

$$(1-\epsilon) \int_X f \, \mathrm{d}\mu \leq \int_X g \, \mathrm{d}\mu \leq \lim_{j \to \infty} \int_X f_j \, \mathrm{d}\mu \leq \int_X f \, \mathrm{d}\mu,$$

which gives

$$\lim_{j \to \infty} \int_X f_j \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu$$

since ϵ is arbitrary. ■

The following corollary to the preceding result ensures consistency of Definition 5.7.4 with Definition 5.7.2.

5.7.6 Corollary (Consistency of integral definitions) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a measure space and if* $\mathsf{f} \in \mathsf{S}(\mathsf{X}; \overline{\mathbb{R}}_{\geq 0})$ *then the integral of* f *as in Definition 5.7.4 agrees with the integral of* f *as in Definition 5.7.2.*

Proof Consider the constant sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ defined by $f_j = f$, $j \in \mathbb{Z}_{>0}$. By Proposition 5.7.5 it follows that the integral of f from Definition 5.7.4 satisfies

$$\int_X f \, \mathrm{d}\mu = \lim_{j \to \infty} \int_X f_j \, \mathrm{d}\mu,$$

where the integrals on the left are as in Definition 5.7.2. However, each of these integrals is exactly the integral of f itself as in Definition 5.7.2. ■

Let us give a somewhat simple application of the preceding result that uses the counting measure. This example is interesting in and of itself as it begins the casting of the notion of summation using general index sets from Section 2.4.7 in the framework of integration on measure spaces; this programme is completed in Example 5.7.10 below. For other examples of integration we shall wait until Sections 5.9 and **??**.

5.7.7 Example (Sums as integrals) Let X be a set, take $\mathscr{A} = \mathbf{2}^X$, and let μ_Σ be the counting measure; see Example 5.3.9–3. Note that all functions $f \colon X \to \overline{\mathbb{R}}$ are measurable. Let $f \in \mathsf{L}^{(0)}((X, \mathscr{A}), \overline{\mathbb{R}}_{\geq 0})$ be a positive nonnegative-valued function. Let us attempt to understand the integral of f. We denote

$$\operatorname{supp}(f) = \{x \in X \mid f(x) \neq 0\}$$

and then consider three cases.

1. $\operatorname{supp}(f)$ *is finite:* Here f is a simple function and we immediately have

$$\int_X f \, \mathrm{d}\mu_\Sigma = \sum_{x \in \operatorname{supp}(f)} f(x),$$

using the definition of the integral of a simple function and the definition of the counting measure.

2. supp(f) *is countably infinite:* In this case we write supp$(f) = \{x_j\}_{j\in\mathbb{Z}_{>0}}$ for distinct $x_j \in X$, $j \in \mathbb{Z}_{>0}$. Let us then define a sequence $(f_k)_{k\in\mathbb{Z}_{>0}}$ of $\overline{\mathbb{R}}_{\geq 0}$-valued functions on X by
$$f_k(x) = \begin{cases} f(x), & x \in \{x_1, \dots, x_k\}, \\ 0, & \text{otherwise.} \end{cases}$$
Then the sequence $(f_k)_{k\in\mathbb{Z}_{>0}}$ is monotonically increasing and satisfies $\lim_{k\to\infty} f_k(x) = f(x)$ for every $x \in X$. Note that the functions f_k, $k \in \mathbb{Z}_{>0}$, are simple and that
$$\int_X f_k \,\mathrm{d}\mu_\Sigma = \sum_{j=1}^{k} f(x_j),$$
using the definition of the integral of a simple function and the definition of the counting measure. Thus, by Proposition 5.7.5 we have
$$\int_X f \,\mathrm{d}\mu_\Sigma = \lim_{k\to\infty} \int_X f_k(x) \,\mathrm{d}\mu_\Sigma = \sum_{j=1}^{\infty} f(x_j).$$
In other words,
$$\int_X f \,\mathrm{d}\mu_\Sigma = \sum_{x\in X} f(x),$$
where the sum is interpreted as in Section 2.4.7, and where we allow the sum to be infinite.

3. supp(f) *is uncountable:* For $k \in \mathbb{Z}_{>0}$ define
$$A_k = \left\{x \in X \;\middle|\; f(x) \geq \tfrac{1}{k}\right\}.$$
We claim that one of the sets A_k must be infinite for some $k \in \mathbb{Z}_{>0}$. Indeed, if all of the sets A_k, $k \in \mathbb{Z}_{>0}$, is finite then, since supp$(f) = \cup_{k\in\mathbb{Z}_{>0}} A_k$, it follows that supp$(f)$ is countable by Proposition **??**. Thus it must be the case that A_k is infinite for some $k \in \mathbb{Z}_{>0}$. In case A_k is uncountable, let A'_k be a countable subset of A_k. Now define $f_k \colon X \to \overline{\mathbb{R}}_{\geq 0}$ by
$$f_k(x) = \begin{cases} f(x), & x \in A'_k, \\ 0, & \text{otherwise.} \end{cases}$$
Note that $f_k(x) \leq f(x)$ for every $x \in X$. Then, using Exercise 5.7.1 and the fact that we know how to integrate f_k from the preceding case, we have
$$\int_X f \,\mathrm{d}\mu_\Sigma \geq \int_{A_k} f \,\mathrm{d}\mu_\Sigma = \sum_{x\in A_k} f_k(x) \geq \sum_{x\in A_k} \frac{1}{k} = \infty.$$
Thus the integral of f is infinite.

Thus, in summary, we have

$$\int_X f \, \mathrm{d}\mu_\Sigma = \sum_{x \in X} f(x),$$

using the definition of series using arbitrary index sets in Section 2.4.7, and with the convention that the integral is allow to be infinite, and indeed will be infinite if $\operatorname{supp}(f)$ is uncountable. •

Integration of general measurable functions

It is now relatively easy to define the integral for general measurable functions on a measure space $(X, \mathscr{A}, \mu)$. To do so, if $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$ we define $f_+, f_- \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ by

$$f_+(x) = \max\{f(x), 0\}, \qquad f_-(x) = \max\{-f(x), 0\},$$

noting that these functions are indeed measurable by Corollary 5.6.17. We may now directly give the definition of the integral.

5.7.8 Definition (Integral of measurable function) For a measure space $(X, \mathscr{A}, \mu)$ and for $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$, we have the following definitions.

(i) If at least one of $\int_X f_+ \, \mathrm{d}\mu$ or $\int_X f_- \, \mathrm{d}\mu$ are finite then the integral of f with respect to μ ***exists*** and is given by

$$\int_X f \, \mathrm{d}\mu = \int_X f_+ \, \mathrm{d}\mu - \int_X f_- \, \mathrm{d}\mu,$$

this being the ***integral*** of f with respect to μ.

(ii) If both $\int_X f_+ \, \mathrm{d}\mu$ and $\int_X f_- \, \mathrm{d}\mu$ are infinite then the integral of f with respect to μ ***does not exist***.

(iii) If $\int_X f_+ \, \mathrm{d}\mu < \infty$ and $\int_X f_- \, \mathrm{d}\mu < \infty$ then f is ***integrable*** with respect to μ.

For a subset $I \subseteq \overline{\mathbb{R}}$ we denote the set of I-valued functions integrable with respect to μ by $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); I)$, or simply by $\mathsf{L}^{(1)}(X; I)$ if $\mathscr{A}$ and μ are understood. •

5.7.9 Notation ($\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$) The notation $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ seems a little odd at this point. For example, what does the superscript "1" mean? And why are there parentheses around the "1." This will be presented in context in Section 6.7.8, so the reader should perhaps not worry at this point what is the precise meaning of the "1." We might mention, however, that the "L" refers to "Lebesgue," as this notation was first used in the context of the Lebesgue integral, and this will be the setting where the notation will be mainly used by us in these volumes. •

Again, we delay until Sections 5.9 and **??** the presentation of examples related to the Lebesgue measure. However, we can at this point complete our example of how the integral includes the usual notion of series.

5.7.10 Example (Sums as integrals (cont'd)) As in Example 5.7.7 we consider a set X, we let $\mathscr{A} = \mathbf{2}^X$, and we let μ_Σ be the counting measure defined in Example 5.3.9–3. We let $f\colon X \to \overline{\mathbb{R}}$, noting again that all functions are measurable. We then note that, as in Example 5.7.7, we have

$$\int_X f_+ \,\mathrm{d}\mu_\Sigma = \sum_{x\in X} f_+(x), \qquad \int_X f_- \,\mathrm{d}\mu_\Sigma = \sum_{x\in X} f_-(x), \tag{5.15}$$

using the notion of sums with arbitrary index sets from Section 2.4.7, and allowing that these quantities may be infinite. Note that the general summation construction of Section 2.4.7, along with the definition of the integral, then immediately gives

$$\int_X f \,\mathrm{d}\mu_\Sigma = \sum_{x\in X} f(x)$$

if either of the sums in (5.15) is finite, and otherwise the integral is undefined.

Using Proposition 2.4.32 we see that in the case that $X = \mathbb{Z}_{>0}$, a function is integrable if and only if the sum $\sum_{j=1}^\infty f(j)$ is absolutely convergent. In this case, the value of the integral is exactly the sum of the series. Thus we see that the construction of the integral we give generalises the notion of an *absolutely* convergent series. Note that it does not generalise the notion of a convergent series. It can be made to do so by using special constructions. We do this for the Lebesgue integral in Sections 5.9.2 and **??**. •

Let us close this section by giving a few more or less obvious properties of the integral.

5.7.11 Proposition (Integrals of functions agreeing almost everywhere) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *have the property that* $\mathsf{f}(\mathsf{x}) = \mathsf{g}(\mathsf{x})$ *for almost every* $\mathsf{x} \in \mathsf{X}$. *Then the integral of* f *exists if and only if the integral of* g *exists, and if either integral exists then we have*

$$\int_\mathsf{X} \mathsf{f}\,\mathrm{d}\mu = \int_\mathsf{X} \mathsf{g}\,\mathrm{d}\mu.$$

Proof By breaking both f and g into their positive and negative parts, we can without loss of generality suppose that both functions take values in $\overline{\mathbb{R}}_{\geq 0}$. Let Z be the set where f and g are not equal and let h take the value ∞ on Z and zero elsewhere. Since $f \leq g + h$ we have

$$\int_X f\,\mathrm{d}\mu \leq \int_X g\,\mathrm{d}\mu + \int_X h\,\mathrm{d}\mu,$$

by Propositions 5.7.16 and 5.7.19. The argument can be reversed to give

$$\int_X g\,\mathrm{d}\mu \leq \int_X f\,\mathrm{d}\mu + \int_X h\,\mathrm{d}\mu,$$

and the result follows since $\int_X h\,\mathrm{d}\mu = 0$. ■

The following simple result comes up on occasion in our presentation, so we state it explicitly. Since the result is "obvious," we shall often use it without mention.

5.7.12 Proposition (Integrable functions are almost everywhere finite) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a measure space and if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *then*

$$\mu\Big(\{\mathsf{x} \in \mathsf{X} \mid \ \mathsf{f}(\mathsf{x}) \notin \mathbb{R}\}\Big) = 0.$$

Proof Since f is integrable if both its positive and negative parts, f_+ and f_-, are integrable, we may as well assume that f takes values in $\overline{\mathbb{R}}_{\geq 0}$. Suppose that $f(x) = \infty$ for $x \in A$ with $\mu(A) > 0$. For $N \in \mathbb{Z}_{>0}$ consider the simple function

$$g_N(x) = \begin{cases} N, & x \in A, \\ 0, & \text{otherwise.} \end{cases}$$

We have $g_N(x) \leq f(x)$ for all $x \in A$ and $\int_X g_N \, \mathrm{d}\mu = N\mu(A) > 0$. By the definition of the integral we have $\int_X f \, \mathrm{d}\mu \geq N\mu(A)$, so showing that the integral of f is not finite, since this holds for all $N \in \mathbb{Z}_{>0}$. ■

5.7.13 Remark (Integrable functions may as well be $\mathbb{R}$-valued) Combining Propositions 5.7.11 and 5.7.12 we see that if $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ then, for the purposes of integration, we may as well suppose that f is $\mathbb{R}$-valued. Indeed, if we define

$$g(x) = \begin{cases} f(x), & f(x) \in \mathbb{R}, \\ 0, & f(x) \in \{-\infty, \infty\}, \end{cases}$$

then $\int_X g \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu$. For this reason, when we discuss spaces of integrable functions in Section 6.7, we will assume all functions are finite-valued. It is really only useful to allow functions to take infinite values when doing constructions with pointwise limits. •

The following result is another "obvious" result that we will use without mention throughout the text.

5.7.14 Proposition (Positive functions with zero integral) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a measure space and if* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ *satisfies* $\int_{\mathsf{X}} \mathsf{f} \, \mathrm{d}\mu = 0$ *then*

$$\mu\Big(\{\mathsf{x} \in \mathsf{X} \mid \ \mathsf{f}(\mathsf{x}) \neq 0\}\Big) = 0.$$

Proof Suppose that $A \subseteq X$ has positive Lebesgue measure and that $f(x) > 0$ for all $x \in A$. Since $f \geq f\chi_A$, by Proposition 5.7.19 it follows that

$$\int_X f \, \mathrm{d}\mu \geq \int_X f\chi_A \, \mathrm{d}\mu > 0,$$

which gives the result. ■

As a final result in this section we record the relationship between functions that are measurable on the completion of a measure space and those that are measurable on the incomplete measure space.

5.7.15 Proposition (Integrable functions on the completion) *Let $(\mathsf{X}, \mathscr{A}_\mu, \overline{\mu})$ be the completion of the measure space $(\mathsf{X}, \mathscr{A}, \mu)$ and let $\mathsf{f}\colon \mathsf{X} \to \overline{\mathbb{R}}$ be $\mathscr{A}_\mu$-measurable. Then there exists a function $\mathsf{g}\colon \mathsf{X} \to \overline{\mathbb{R}}$ that is $\mathscr{A}$-measurable and with the property that*

$$\mu(\{\mathsf{x} \in \mathsf{X} \mid \mathsf{g}(\mathsf{x}) \neq \mathsf{f}(\mathsf{x})\}) = 0.$$

Moreover, the integral of f with respect to $\overline{\mu}$ exists if and only if the integral of g with respect to μ exists, and in this case,

$$\int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\overline{\mu} = \int_{\mathsf{X}} \mathsf{g}\, \mathrm{d}\mu.$$

Proof First suppose that f takes values in $\overline{\mathbb{R}}_{\geq 0}$. By Proposition 5.6.39 let $(g_j)_{j \in \mathbb{Z}_{>0}}$ be a monotonically increasing sequence of simple functions for which $\lim_{j\to\infty} g_j(x) = f(x)$ for all $x \in X$. This means that we may write f as an infinite sum of characteristic functions:

$$f(x) = \sum_{j=1}^{\infty} c_j \chi_{A_j}(x),$$

where $c_j \in \mathbb{R}_{\geq 0}$ and $A_j \in \mathscr{A}_\mu$, $j \in \mathbb{Z}_{>0}$. For $j \in \mathbb{Z}_{>0}$ let $L_j, U_j \in \mathscr{A}$ have the property that $L_j \subseteq A_j \subseteq U_j$ and $\mu(U_j \setminus L_j) = 0$. Taking

$$g(x) = \sum_{j=1}^{\infty} c_j \chi_{U_j}(x)$$

for $x \in X$ gives the first part of the result in this case since f and g differ on the set $(\cup_{j \in \mathbb{Z}_{>0}} U_j \setminus A_j) \subseteq (\cup_{j \in \mathbb{Z}_{>0}} U_j \setminus L_j)$, and this latter set has measure zero by Exercise 5.3.4.

Now suppose that f is now allowed to take arbitrary values in $\overline{\mathbb{R}}$. Write $f = f_+ - f_-$, where

$$f_+(x) = \max\{f(x), 0\}, \qquad f_-(x) = \max\{-f(x), 0\}.$$

These functions are $\mathscr{A}_\mu$-measurable by Corollary 5.6.17. Therefore, there exist $\mathscr{A}$-measurable functions g_+ and g_- such that f_+ differs from g_+ and f_- differs from g_- on a set of measure zero. Therefore, f differs from $g = g_+ - g_-$ on a set of measure zero. The result follows since g is $\mathscr{A}$-measurable by Proposition 5.6.11.

Now let us prove the last assertion of the proposition. Write $f = g + h$ for f being $\mathscr{A}_\mu$-measurable, for g being $\mathscr{A}$-measurable, and for

$$\mu(\{x \in \mathrm{X} \mid h(x) \neq 0\}) = 0.$$

Let $Z \in \mathscr{A}$ be a set such that $h(x) = 0$ for $x \in \mathrm{X} \setminus Z$ and such that $\mu(Z) = 0$. Then

$$\int_X f\, \mathrm{d}\overline{\mu} = \int_{X\setminus Z} g\, \mathrm{d}\overline{\mu} + \int_Z (g+h)\, \mathrm{d}\overline{\mu} = \int_{X\setminus Z} g\, \mathrm{d}\overline{\mu} = \int_X g\, \mathrm{d}\overline{\mu},$$

using Proposition 5.7.11. Now note that since g if integrable with respect to μ, its integral with respect to $\overline{\mu}$ can be constructed using the definition of the integral without reference to the distinction between μ and $\overline{\mu}$. That is to say,

$$\int_X g\, \mathrm{d}\overline{\mu} = \int_X g\, \mathrm{d}\mu,$$

and from this the result follows. ■

5.7.2 The integral and operations on functions

In this section we provide the more or less expected result regarding the interaction of the integral with the standard operations one may perform on functions. It is useful to record two different versions of results, one for arbitrary positive measurable functions and one for integrable functions.

We begin with the relationships between the integral and the standard algebraic operations on functions. We recall from Proposition 5.6.11 that $\mathsf{L}^{(0)}((X,\mathscr{A});\overline{\mathbb{R}})$ is a subset of the set $\overline{\mathbb{R}}^X$ of all $\overline{\mathbb{R}}$-valued functions on X, and this subset is closed under addition and multiplication on $\overline{\mathbb{R}}$. With this in mind we have the following results.

5.7.16 Proposition (Algebraic operations on positive measurable functions) *For a measure space* $(\mathsf{X},\mathscr{A},\mu)$*, for* $\mathsf{f},\mathsf{g}\in\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A});\overline{\mathbb{R}}_{\geq 0})$*, and for* $\alpha\in\overline{\mathbb{R}}_{\geq 0}$*, the following statements hold:*

(i) $\displaystyle\int_{\mathsf{X}}(\mathsf{f}+\mathsf{g})\,\mathrm{d}\mu=\int_{\mathsf{X}}\mathsf{f}\,\mathrm{d}\mu+\int_{\mathsf{X}}\mathsf{g}\,\mathrm{d}\mu;$

(ii) $\displaystyle\int_{\mathsf{X}}\alpha\mathsf{f}\,\mathrm{d}\mu=\alpha\int_{\mathsf{X}}\mathsf{f}\,\mathrm{d}\mu.$

Proof We let $(f_j)_{j\in\mathbb{Z}_{>0}}$ and $(g_j)_{j\in\mathbb{Z}_{>0}}$ be sequences of simple functions converging to f and g, respectively, as in Proposition 5.6.39.

(i) Note that if either $\lim_{j\to\infty}f_j(x)$ or $\lim_{j\to\infty}g_j(x)$ is infinite, then

$$\lim_{j\to\infty}(f_j+g_j)(x)=\lim_{j\to\infty}f_j(x)+\lim_{j\to\infty}g_j(x)=f(x)+g(x)=\infty.$$

If both $\lim_{j\to\infty}f_j(x)$ and $\lim_{j\to\infty}g_j(x)$ are finite then we have

$$\lim_{j\to\infty}(f_j+g_j)(x)=\lim_{j\to\infty}f_j(x)+\lim_{j\to\infty}g_j(x)=f(x)+g(x)$$

by Proposition 2.3.23. Thus $(f_j+g_j)_{j\in\mathbb{Z}_{>0}}$ is a monotonically increasing sequence of simple functions converging to $f+g$. Thus this part of the result will follow from Proposition 5.7.5 if we can establish it for simple functions. Thus we assume that f and g are simple functions and denote

$$f=\sum_{j=1}^{k}a_j\chi_{A_j},\quad g=\sum_{l=1}^{m}b_l\chi_{B_l}.$$

for $a_1,\dots,a_k,b_1,\dots,b_l\in\overline{\mathbb{R}}$ and $A_1,\dots,A_k$ and $B_1,\dots,B_m$ are disjoint. We assume without loss of generality that $\cup_{j=1}^{k} A_j = \cup_{l=1}^{m} B_l$. Then

$$\begin{aligned}\int_A (f+g)\,\mathrm{d}\mu &= \sum_{j=1}^{k}\sum_{l=1}^{m}(a_j+b_l)\mu(A_j\cap B_l)\\ &= \sum_{j=1}^{k}\sum_{l=1}^{m} a_j\mu(A_j\cap B_l) + \sum_{j=1}^{k}\sum_{l=1}^{m} b_l\mu(A_j\cap B_l)\\ &= \sum_{j=1}^{k} a_j\mu(A_j) + \sum_{l=1}^{m} b_l\mu(B_l)\\ &= \int_A f\,\mathrm{d}\mu + \int_A g\,\mathrm{d}\mu,\end{aligned}$$

so giving (i).

(ii) If either α or $\lim_{j\to\infty} f_j(x)$ is infinite then obviously we have

$$\lim_{j\to\infty} \alpha f_j(x) = \alpha f(x) = \infty.$$

If both α and $\lim_{j\to\infty} f_j(x)$ are finite then we have

$$\lim_{j\to\infty} \alpha f_j(x) = \alpha f(x)$$

by Proposition 2.3.23. Thus $(\alpha f_j)_{j\in\mathbb{Z}_{>0}}$ is a monotonically increasing sequence of positive simple functions that converges to αf. Part (ii) then follows from Proposition 5.7.5. ■

5.7.17 Proposition (Algebraic operations on integrable functions) *For a measure space* $(\mathsf{X}, \mathscr{A}, \mu)$*, for* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$*, and for* $\alpha \in \mathbb{R}$*, the following statements hold:*

(i) $\mathsf{f} + \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and*

$$\int_{\mathsf{X}} (\mathsf{f}+\mathsf{g})\,\mathrm{d}\mu = \int_{\mathsf{X}} \mathsf{f}\,\mathrm{d}\mu + \int_{\mathsf{X}} \mathsf{g}\,\mathrm{d}\mu;$$

(ii) $\alpha\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and*

$$\int_{\mathsf{X}} \alpha\mathsf{f}\,\mathrm{d}\mu = \alpha\int_{\mathsf{X}} \mathsf{f}\,\mathrm{d}\mu.$$

Proof The proposition follows from Proposition 5.7.16 by breaking f and g into their positive and negative parts, and applying the lemma to both resulting integrals. ■

One might wonder about the relationships between integrals and other algebraic operations on functions, like multiplication and division. Generally speaking, these operations fail to preserve integrability.

5.7.18 Examples (Multiplication, division, and the integral)

1. We take $X = \mathbb{Z}_{>0}$ with the σ-algebra $\mathscr{A} = \mathbf{2}^{\mathbb{Z}_{>0}}$ and the counting measure μ_Σ. In this case, integrable functions are those functions $f\colon \mathbb{Z}_{>0} \to \mathbb{R}$ satisfying $\sum_{j=1}^{\infty}|f(j)| < \infty$; this follows from Example 5.7.10, or more straightforwardly from Exercise 5.7.3. Let us define $f\colon \mathbb{Z}_{>0} \to \mathbb{R}$ by $f(j) = \frac{1}{j^2}$. By Example 2.4.2–**??** it follows that $f \in \mathsf{L}^{(1)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}, \mu_\Sigma); \mathbb{R})$. However, since $f^2(j) = \frac{1}{j}$, it follows from Example 2.4.2–**??** that $f^2 \notin \mathsf{L}^{(1)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}, \mu_\Sigma); \mathbb{R})$. Thus products of integrable functions need not be integrable functions.
2. We take $X = \mathbb{Z}_{>0}$, $\mathscr{A} = \mathbf{2}^{\mathbb{Z}_{>0}}$, and $\mu = \mu_\Sigma$ as in the previous example. We note that if we define $f, g\colon \mathbb{Z}_{>0} \to \mathbb{R}$ by $f(j) = \frac{1}{j^2}$ and $g(j) = \frac{1}{j^3}$; as above, $f, g \in \mathsf{L}^{(1)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}, \mu_\Sigma); \mathbb{R})$. However, clearly $\frac{f}{g}(j) = \frac{1}{j}$ and so $\frac{f}{g} \notin \mathsf{L}^{(1)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}, \mu_\Sigma); \mathbb{R})$. Thus the quotient of two integrable functions is not necessarily integrable, even when the denominator function is nowhere zero. •

For functions whose values are related by the total order on $\overline{\mathbb{R}}$ we have the following result applies.

5.7.19 Proposition (The integral and total order on $\overline{\mathbb{R}}$) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a measure space and if* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *(resp.* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$*) satisfy* $\mathsf{f}(\mathsf{x}) \leq \mathsf{g}(\mathsf{x})$ *for almost all* $\mathsf{x} \in \mathsf{X}$*, then*

$$\int_{\mathsf{X}} \mathsf{f}\,\mathrm{d}\mu \leq \int_{\mathsf{X}} \mathsf{g}\,\mathrm{d}\mu.$$

Proof Without loss of generality we may suppose that $f(x) \leq g(x)$ for all $x \in X$. Indeed, if this inequality holds except on a set Z which has zero measure, then we have

$$\int_X f\,\mathrm{d}\mu = \int_{X\setminus Z} f\,\mathrm{d}\mu + \int_Z f\,\mathrm{d}\mu = \int_{X\setminus Z} f\,\mathrm{d}\mu,$$

and so we can simply replace X with $X \setminus Z$.

Now we may use part (i) from Proposition 5.7.16 or Proposition 5.7.17 to write

$$\int_X g\,\mathrm{d}\mu = \int_X (f + (g - f))\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu + \int_X (g - f)\,\mathrm{d}\mu \geq \int_A f\,\mathrm{d}\mu,$$

as desired. ■

This result has the following corollary which we often apply.

5.7.20 Corollary (Functions bounded by integrable functions are integrable) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *satisfy* $|\mathsf{f}(\mathsf{x})| \leq |\mathsf{g}(\mathsf{x})|$ *for almost every* $\mathsf{x} \in \mathsf{X}$*. If* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *then* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$.

Proof Write $f = f_+ - f_-$ and $g = g_+ - g_-$ for $f_+, g_+, f_-, g_- \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$. Then we obviously have

$$f_+(x) \leq g_+(x), \quad f_- \leq g_-(x)$$

for almost every $x \in X$. Thus, by Proposition 5.7.19 we have

$$\int_X f_+ \, d\mu \leq \int_X g_+ \, d\mu, \quad \int_X f_- \, d\mu \leq \int_X g_- \, d\mu.$$

Therefore,

$$\int_X |f| \, d\mu = \int_X f_+ \, d\mu + \int_X f_- \, d\mu \leq \int_X g_+ \, d\mu + \int_X g_- \, d\mu = \int_X |g| \, d\mu,$$

as desired. ■

The following result follows pretty much from the definitions surrounding the Lebesgue integral.

5.7.21 Proposition (The integral and absolute value) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$. *Then* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $|\mathsf{f}| \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$, *and if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *then*

$$\left| \int_{\mathsf{X}} \mathsf{f} \, d\mu \right| \leq \int_{\mathsf{X}} |\mathsf{f}| \, d\mu.$$

Proof The first assertion is Exercise 5.7.4. For the second assertion, write $f = f_+ - f_-$ as the sum of its positive and negative parts. Then

$$\left| \int_X f \, d\mu \right| \leq \left| \int_X f_+ \, d\mu \right| + \left| \int_X f_- \, d\mu \right| = \int_X |f| \, d\mu,$$

using the fact that for a positive function the integral is positive. ■

It is at times useful to break an integral into two parts by breaking the domain of integration into two parts.

5.7.22 Proposition (Breaking the integral in two) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space, let* $\mathsf{A}, \mathsf{B} \in \mathscr{A}$ *be sets such that* $\mathsf{X} = \mathsf{A} \mathbin{\overset{\circ}{\cup}} \mathsf{B}$, *and let* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$. *Then* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $\mathsf{f}|\mathsf{A} \in \mathsf{L}^{(1)}((\mathsf{A}, \mathscr{A}_{\mathsf{A}}, \mu|\mathscr{A}_{\mathsf{A}}); \overline{\mathbb{R}})$ *and* $\mathsf{f}|\mathsf{B} \in \mathsf{L}^{(1)}((\mathsf{B}, \mathscr{A}_{\mathsf{b}}, \mu|\mathscr{A}_{\mathsf{B}}); \overline{\mathbb{R}})$. *Furthermore, if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *then we have*

$$\int_{\mathsf{X}} \mathsf{f} \, d\mu = \int_{\mathsf{A}} (\mathsf{f}|\mathsf{A}) \, d\mu_{\mathsf{A}} + \int_{\mathsf{B}} (\mathsf{f}|\mathsf{B}) \, d\mu_{\mathsf{B}}.$$

Proof Let us define $f_A, f_B \colon X \to \overline{\mathbb{R}}$ by $f_A = f\chi_A$ and $f_B = f\chi_B$. By Proposition 5.6.15 the functions f_A and f_B are measurable. We claim that, provided that f_A and $f|A$ are integrable,

$$\int_X f_A \, d\mu = \int_A (f|A) \, d\mu_A. \tag{5.16}$$

To see this, first suppose that f is $\overline{\mathbb{R}}_{\geq 0}$-valued and let $(f_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence of simple functions converging to f as in Proposition 5.6.39. Then the sequence $(f_{A,j})_{j \in \mathbb{Z}_{>0}}$ defined by $f_{A,j} = f_j \chi_A$ is a sequence of simple functions converging to f_A as in Proposition 5.6.39. Moreover,

$$\int_X f_{A,j} \, d\mu = \int_A (f_j|A) \, d\mu_A$$

by Exercise 5.7.2. Therefore, by Proposition 5.7.5 we have

$$\int_X f_A \,\mathrm{d}\mu = \lim_{j\to\infty} \int_X f_{A,j} \,\mathrm{d}\mu = \lim_{j\to\infty} \int_A (f_j|A) \,\mathrm{d}\mu_A = \int_A (f|A) \,\mathrm{d}\mu_A,$$

giving (5.16) when f is $\overline{\mathbb{R}}_{\geq 0}$-valued. For $\overline{\mathbb{R}}$-valued f the same conclusion follows by breaking f into its positive and negative parts. Similarly, of course, we have

$$\int_X f_B \,\mathrm{d}\mu = \int_B (f|B) \,\mathrm{d}\mu_B,$$

and so Proposition 5.7.17 gives the final assertion of the result provided that f, f_A, and f_B are integrable.

Now, if f_A and f_B are integrable, by Proposition 5.7.17 it follows that f is integrable. Conversely, if either of f_A or f_B are not integrable, then neither can f be integrable (why?). ■

A more general version of the preceding result is useful, but is only valid for complete measure spaces.

5.7.23 Corollary (Breaking the integral almost in two) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a complete measure space, let* $\mathsf{A}, \mathsf{B} \in \mathscr{A}$ *be such that* $\mu(\mathsf{A} \cap \mathsf{B}) = 0$ *and such that* $\mathsf{X} = \mathsf{A} \cup \mathsf{B}$*, and let* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*. Then* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $\mathsf{f}|\mathsf{A} \in \mathsf{L}^{(1)}((\mathsf{A}, \mathscr{A}_{\mathsf{A}}, \mu|\mathscr{A}_{\mathsf{A}}); \overline{\mathbb{R}})$ *and* $\mathsf{f}|\mathsf{B} \in \mathsf{L}^{(1)}((\mathsf{B}, \mathscr{A}_{\mathsf{B}}, \mu|\mathscr{A}_{\mathsf{B}}); \overline{\mathbb{R}})$*. Furthermore, if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *then we have*

$$\int_{\mathsf{X}} \mathsf{f} \,\mathrm{d}\mu = \int_{\mathsf{A}} (\mathsf{f}|\mathsf{A}) \,\mathrm{d}\mu_{\mathsf{A}} + \int_{\mathsf{B}} (\mathsf{f}|\mathsf{B}) \,\mathrm{d}\mu_{\mathsf{B}}.$$

Proof Let $Z = A \cap B$, let $A' = A - Z$ and $B' = B - Z$ and write $X = A' \overset{\circ}{\cup} B' \overset{\circ}{\cup} Z$. Note that Z, A', and B' are measurable since X is complete. Applying Proposition 5.7.22 (or more properly, its obvious extension to finitely many disjoint components) gives

$$\int_X f \,\mathrm{d}\mu = \int_{A'} (f|A') \,\mathrm{d}\mu_{A'} + \int_{B'} (f|B')\mathrm{d}\mu_{B'} + \int_Z (f|Z) \,\mathrm{d}\mu.$$

The last integral is zero by Proposition 5.7.11 and, by the same result,

$$\int_{A'} (f|A') \,\mathrm{d}\mu_{A'} = \int_A (f|A) \,\mathrm{d}\mu_A$$

and

$$\int_{B'} (f|B') \,\mathrm{d}\mu_{B'} = \int_B (f|B) \,\mathrm{d}\mu_B,$$

giving the result. ■

5.7.3 Limit theorems

In Section 5.1 we suggested that one of the reasons why the Riemann integral was not satisfactory was that it did not have useful properties with respect to swapping of limits and integration. In this section we prove some powerful theorems for the integral on measure spaces which give very general conditions under which limits and integrals will swap. When these are applied to the Lebesgue integral in Sections 5.9 and **??**, this will show that we have produced a theory of integration that generalises the Riemann integral, and which has at least some more desirable properties.

Our first theorem has very weak hypotheses, but only applies to nonnegative functions.

5.7.24 **Theorem (Monotone Convergence Theorem I)** *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $(\mathrm{f}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ *such that, for almost every* $\mathrm{x} \in \mathsf{X}$, $\mathrm{f}_\mathrm{j}(\mathrm{x}) \leq \mathrm{f}_{\mathrm{j}+1}(\mathrm{x})$ *for every* $\mathrm{j} \in \mathbb{Z}_{>0}$*. If* $\mathrm{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *has the property that* $\mathrm{f}(\mathrm{x}) = \lim_{\mathrm{j}\to\infty} \mathrm{f}_\mathrm{j}(\mathrm{x})$ *for almost every* $\mathrm{x} \in \mathsf{X}$*, then*

$$\int_{\mathsf{X}} \mathrm{f}\, \mathrm{d}\mu = \lim_{\mathrm{j}\to\infty} \int_{\mathsf{X}} \mathrm{f}_\mathrm{j}\, \mathrm{d}\mu.$$

Proof First let us show that we may assume without loss of generality that the relations $f_j(x) \leq f_{j+1}(x)$, $j \in \mathbb{Z}_{>0}$, and $f(x) = \lim_{j\to\infty} f_j(x)$ hold for all $x \in X$. Let Z be the set on which these relation do not hold, noting that Z has measure zero being a union of two sets of measure zero. Let $Y = X \setminus Z$. The sequence of functions $(f_j\chi_Y)_{j\in\mathbb{Z}_{>0}}$ and the function $f\chi_Y$ then satisfy the relations for all $x \in X$. If the theorem holds in this case, then the result will follow from Proposition 5.7.11. For the remainder of the proof we therefore assume that $f_j(x) \leq f_{j+1}(x)$, $j \in \mathbb{Z}_{>0}$, and $f(x) = \lim_{j\to\infty} f_j(x)$ for all $x \in X$.

By Proposition 5.7.19 we have

$$\int_X f_j\, \mathrm{d}\mu \leq \int_X f_{j+1}\, \mathrm{d}\mu, \qquad j \in \mathbb{Z}_{>0},$$

and

$$\int_X f_j\, \mathrm{d}\mu \leq \int_X f\, \mathrm{d}\mu, \qquad j \in \mathbb{Z}_{>0}.$$

Thus the sequence $(\int_X f_j\, \mathrm{d}\mu)_{j\in\mathbb{Z}_{>0}}$ converges in $\overline{\mathbb{R}}_{\geq 0}$ to a limit and this limit satisfies

$$\lim_{j\to\infty} \int_X f_j\, \mathrm{d}\mu \leq \int_X f\, \mathrm{d}\mu.$$

We wish to establish the opposite inequality. For each $j \in \mathbb{Z}_{>0}$ let $(g_{j,k})_{k\in\mathbb{Z}_{>0}}$ be a sequence of simple functions whose limit is f_j, as in Proposition 5.6.39. Now define $h_k(x) = \max\{g_{1,k}(x), \dots, g_{k,k}(x)\}$, and note that $(h_k)_{k\in\mathbb{Z}_{>0}}$ is a monotonically increasing sequence of simple functions converging to f, and that $h_k(x) \leq f_k(x)$ for all $x \in X$. By our above arguments for the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$, we have

$$\lim_{k\to\infty} \int_X h_k\, \mathrm{d}\mu \leq \lim_{j\to\infty} \int_X f_j\, \mathrm{d}\mu \leq \int_X f\, \mathrm{d}\mu.$$

The theorem now follows by Proposition 5.7.5. ■

In the next assertion, the condition that the functions be nonnegative is relaxed, but one must add an integrability condition for one of the functions in the sequence.

5.7.25 Theorem (Monotone Convergence Theorem II) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $(\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *such that, for almost every* $\mathsf{x} \in \mathsf{X}$, $\mathsf{f}_j(\mathsf{x}) \leq \mathsf{f}_{j+1}(\mathsf{x})$ *for every* $j \in \mathbb{Z}_{>0}$ *and such that* $\mathsf{f}_1 \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$*. If* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ *has the property that* $\mathsf{f}(\mathsf{x}) = \lim_{j\to\infty} \mathsf{f}_j(\mathsf{x})$ *for almost every* $\mathsf{x} \in \mathsf{X}$*, then*

$$\int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\mu = \lim_{j\to\infty} \int_{\mathsf{X}} \mathsf{f}_j\, \mathrm{d}\mu.$$

Proof As in the proof of Theorem 5.7.24 we can assume that $f_j(x) \leq f_{j+1}(x)$, $j \in \mathbb{Z}_{>0}$, and $f(x) = \lim_{j\to\infty} f_j(x)$ for every $x \in X$. Note that the sequence $(f_j - f_1)$ is then in $\mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ and satisfies $\lim_{j\to\infty}(f_j(x) - f_1(x)) = f(x) - f_1(x)$ for every $x \in X$. Note that if $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ then we have

$$\int_X f\, \mathrm{d}\mu - \int_X f_1\, \mathrm{d}\mu = \int_X (f - f_1)\, \mathrm{d}\mu$$

by Proposition 5.7.17. If $f \notin \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ then the previous relation still holds with value ∞ on both sides (why?). Therefore, by Theorem 5.7.24, we have

$$\int_X f\, \mathrm{d}\mu - \int_X f_1\, \mathrm{d}\mu = \int_X (f - f_1)\, \mathrm{d}\mu = \lim_{j\to\infty} \int_X (f_j - f_1)\, \mathrm{d}\mu = \lim_{j\to\infty} \int_X f_j\, \mathrm{d}\mu - \int_X f_1\, \mathrm{d}\mu,$$

which gives the result since $f_1 \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$. ■

We also have the following immediate corollary to the Monotone Convergence Theorem.

5.7.26 Corollary (Beppo Levi's[9] Theorem) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $(\mathsf{f}_j)_{j\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$*. If* $\mathsf{f}\colon \mathsf{X} \to \overline{\mathbb{R}}_{\geq 0}$ *is defined by*

$$\mathsf{f}(\mathsf{x}) = \sum_{j=1}^{\infty} \mathsf{f}_j(\mathsf{x}),$$

then f *is measurable and we have*

$$\int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\mu = \sum_{j=1}^{\infty} \int_{\mathsf{X}} \mathsf{f}_j\, \mathrm{d}\mu.$$

Proof Define $g_k(x) = \sum_{j=1}^{k} f_j(x)$, noting that $g_k \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$ by Proposition 5.6.11. Moreover, for every $x \in X$ we have $g_k(x) \leq g_{k+1}(x)$. Thus Theorem 5.7.24 and Proposition 5.7.16 imply that

$$\int_X f\, \mathrm{d}\mu = \lim_{k\to\infty} \int_X g_k\, \mathrm{d}\mu = \lim_{k\to\infty} \sum_{j=1}^{k} \int_X f_j\, \mathrm{d}\mu = \sum_{j=1}^{\infty} \int_X f_j\, \mathrm{d}\mu,$$

as desired. ■

The following result is also useful, but with weaker hypotheses and conclusions than the Monotone Convergence Theorem.

[9] Beppo Levi (1875–1961) was an Italian mathematician who made mathematical contributions to algebra and analysis. As a Jew, he left Italy after the rise of Mussolini for Argentina, where he spent much of his professional life.

5.7.27 Theorem (Fatou's[10] Lemma) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a measure space and if* $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *is a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}}_{\geq 0})$*, then*

$$\int_\mathsf{X} \liminf_{\mathsf{j}\to\infty} \mathsf{f}_\mathsf{j} \, \mathrm{d}\mu \leq \liminf_{\mathsf{j}\to\infty} \int_\mathsf{X} \mathsf{f}_\mathsf{j} \, \mathrm{d}\mu.$$

Proof For $k \in \mathbb{Z}_{>0}$ define $g_k(x) = \inf_{j\geq k} f_j(x)$, noting that g_k so defined is measurable by Proposition 5.6.18. We then note that the sequence $(g_k)_{k\in\mathbb{Z}_{>0}}$ is increasing and that

$$\liminf_{j\to\infty} f_j(x) = \lim_{k\to\infty} g_k(x)$$

for $x \in X$. From Theorem 5.7.24 we then have

$$\int_X \liminf_{j\to\infty} f_j \, \mathrm{d}\mu = \lim_{k\to\infty} \int_X g_k \, \mathrm{d}\mu \leq \liminf_{j\to\infty} \int_X f_j \, \mathrm{d}\mu,$$

since $g_j(x) \leq f_j(x)$ for $j \in \mathbb{Z}_{>0}$ and $x \in X$. ■

The most frequently useful of the limit theorems is the following. It is a result that is used with great regularity in integration theory. For example, many of the fundamental results we state in Sections 6.7 and 8.3 and in Chapters 12 and 13 rely at their core on this important theorem.

5.7.28 Theorem (Dominated Convergence Theorem I) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *having the following properties:*

(i) the limit $\mathsf{f}(\mathsf{x}) = \lim_{\mathsf{j}\to\infty} \mathsf{f}_\mathsf{j}(\mathsf{x})$ *exists for almost every* $\mathsf{x} \in \mathsf{X}$*;*

(ii) there exists $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}}_{\geq 0})$ *such that, for almost every* $\mathsf{x} \in \mathsf{X}$, $|\mathsf{f}_\mathsf{j}(\mathsf{x})| \leq \mathsf{g}(\mathsf{x})$ *for every* $\mathsf{j} \in \mathbb{Z}_{>0}$.

Then the functions f *and* f_j, $\mathsf{j} \in \mathbb{Z}_{>0}$, *are integrable and*

$$\int_\mathsf{X} \mathsf{f} \, \mathrm{d}\mu = \lim_{\mathsf{j}\to\infty} \int_\mathsf{X} \mathsf{f}_\mathsf{j} \, \mathrm{d}\mu.$$

Proof The integrability of f and f_j, $j \in \mathbb{Z}_{>0}$, follows from Corollary 5.7.20. As with our proof of Theorem 5.7.24, we can without loss of generality suppose that (i) and (ii) hold for all $x \in X$. Furthermore, since g is integrable, we may as well suppose that $g(x) \in \mathbb{R}$ for every x, again by Proposition 5.7.11. The sequence $(g + f_j)_{j\in\mathbb{Z}_{>0}}$ is then a sequence of nonnegative functions for which

$$\lim_{j\to\infty} (g + f_j)(x) = (g + f)(x), \qquad x \in X.$$

By Fatou's Lemma this gives

$$\begin{aligned} &\int_X (g + f) \, \mathrm{d}\mu \leq \liminf_{j\to\infty} \int_X (g + f_j) \, \mathrm{d}\mu \\ \implies \quad &\int_X f \, \mathrm{d}\mu \leq \liminf_{j\to\infty} \int_X f_j \, \mathrm{d}\mu. \end{aligned}$$

[10] Pierre Joseph Louis Fatou (1878–1929) was a French mathematician who made substantial contributions to analysis, particularly complex analysis.

Similarly we can show that

$$\begin{aligned} &\int_X (g-f)\,\mathrm{d}\mu \le \liminf_{j\to\infty} \int_X (g-f_j)\,\mathrm{d}\mu \\ \implies \quad &\int_X f\,\mathrm{d}\mu \le \limsup_{j\to\infty} \int_X f_j\,\mathrm{d}\mu. \end{aligned}$$

This gives the result. ■

The Dominated Convergence Theorem also has the following weaker form for more general sequences.

5.7.29 Theorem (Dominated Convergence Theorem II) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((\mathsf{X},\mathscr{A}); \overline{\mathbb{R}})$ *for which there exists* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X},\mathscr{A},\mu); \overline{\mathbb{R}}_{\ge 0})$ *such that, for almost every* $\mathsf{x} \in \mathsf{X}$, $|\mathsf{f}_\mathsf{j}(\mathsf{x})| \le \mathsf{g}(\mathsf{x})$ *for every* $\mathsf{j} \in \mathbb{Z}_{>0}$*. Then the functions* f_j*,* $\mathsf{j} \in \mathbb{Z}_{>0}$*, are integrable and*

(i) $\displaystyle\int_\mathsf{X} \liminf_{\mathsf{j}\to\infty} \mathsf{f}_\mathsf{j}\,\mathrm{d}\mu \le \liminf_{\mathsf{j}\to\infty} \int_\mathsf{X} \mathsf{f}_\mathsf{j}\,\mathrm{d}\mu$ *and*

(ii) $\displaystyle\int_\mathsf{X} \limsup_{\mathsf{j}\to\infty} \mathsf{f}_\mathsf{j}\,\mathrm{d}\mu \ge \limsup_{\mathsf{j}\to\infty} \int_\mathsf{X} \mathsf{f}_\mathsf{j}\,\mathrm{d}\mu.$

Proof The proofs for both conclusions are similar, so we only prove (i). The integrability f_j, $j \in \mathbb{Z}_{>0}$, follows from Corollary 5.7.20. The measurability of $x \mapsto \liminf_{j\to\infty} f_j(x)$ follows from Proposition 5.6.18. As in the proof of Theorem 5.7.24 we may as well assume that $|f_j(x)| \le |g(x)|$ for all $x \in X$ and $j \in \mathbb{Z}_{>0}$. In this case, the sequence $(g+f_j)_{j\in\mathbb{Z}_{>0}}$ is a sequence in $\mathsf{L}^{(0)}((X,\mathscr{A}); \overline{\mathbb{R}}_{\ge 0})$ and so, by Fatou's Lemma and Proposition 5.7.16, we have

$$\begin{aligned} \int_X g\,\mathrm{d}\mu + \int_X \liminf_{j\in\infty} f_j\,\mathrm{d}\mu &= \int_X \liminf_{j\to\infty}(g+f_j)\,\mathrm{d}\mu \\ &\le \liminf_{j\to\infty} \int_X (g+f_j)\,\mathrm{d}\mu \\ &= \int_X g\,\mathrm{d}\mu + \liminf_{j\to\infty} \int_X f_j\,\mathrm{d}\mu, \end{aligned}$$

which gives the result since $g \in \mathsf{L}^{(1)}((X,\mathscr{A},\mu); \overline{\mathbb{R}})$. ■

Let us illustrate how one might use the preceding results.

5.7.30 Examples (Illustration of limit theorems) In both of the examples, we consider the measure space $(X,\mathscr{A},\mu)$ with $X = \mathbb{Z}_{>0}$, $\mathscr{A} = \mathbf{2}^{\mathbb{Z}_{>0}}$, and $\mu = \mu_\Sigma$, the counting measure. Thus, as we have seen in Example 5.7.10, integrable functions are absolutely convergent series.

1. The Monotone Convergence Theorem is often helpful for showing that a certain integral diverges. Let us illustrate this as follows. We wish to ascertain whether the limit

$$\lim_{\alpha\downarrow 1} \sum_{k=1}^{\infty} \frac{1}{k^\alpha} \tag{5.17}$$

exists. Let us define $f_\alpha \in \mathsf{L}^{(0)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}); \overline{\mathbb{R}}_{\geq 0})$ by $f_\alpha(k) = \frac{1}{k^\alpha}$ for $\alpha \in [1,2]$. Let $(\alpha_j)_{j\in\mathbb{Z}_{>0}}$ be a strictly monotonically decreasing sequence such that $\alpha_1 = 2$ and $\lim_{j\to\infty} \alpha_j = 1$. We then have

$$\lim_{\alpha\downarrow 1} \sum_{k=1}^{\infty} \frac{1}{k^\alpha} = \lim_{\alpha\downarrow 1} \int_{\mathbb{Z}_{>0}} f_\alpha \, \mathrm{d}\mu_\Sigma = \lim_{j\to\infty} \int_{\mathbb{Z}_{>0}} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma.$$

Note that $f_{\alpha_j}(k) < f_{\alpha_{j+1}}(k)$ for every $k \in \mathbb{Z}_{>0}$ and $j \in \mathbb{Z}_{>0}$. Therefore, the sequence $(f_{\alpha_j})_{j\in\mathbb{Z}_{>0}}$ satisfies the hypotheses of the Monotone Convergence Theorem. Therefore, we have

$$\lim_{j\to\infty} \int_{\mathbb{Z}_{>0}} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma = \int_{\mathbb{Z}_{>0}} \lim_{j\to\infty} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma = \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

Thus the limit (5.17) does not exist, at least not in $\mathbb{R}$.

2. Let us use the Dominated Convergence Theorem to determine the value of the following limit:

$$\lim_{\alpha\downarrow 1} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^{2\alpha}}.$$

We proceed much as above, defining $f_\alpha \in \mathsf{L}^{(0)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}); \overline{\mathbb{R}})$ by $f_\alpha(k) = \frac{(-1)^{k+1}}{k^{2\alpha}}$. We let $(\alpha_j)_{j\in\mathbb{Z}_{>0}}$ be a strictly monotonically decreasing sequence such that $\alpha_1 = 2$ and $\lim_{j\to\infty} \alpha_j = 1$. It then holds that

$$\lim_{\alpha\downarrow 1} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^{2\alpha}} = \lim_{\alpha\downarrow 1} \int_{\mathbb{Z}_{>0}} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma = \lim_{j\to\infty} \int_{\mathbb{Z}_{>0}} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma.$$

We then have

$$|f_{\alpha_j}(k)| = \frac{1}{k^{2\alpha_j}} < \frac{1}{k^2}$$

for every $j \in \mathbb{Z}_{>0}$ and $k \in \mathbb{Z}_{>0}$. Define $g \in \mathsf{L}^{(0)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}}); \overline{\mathbb{R}})$ by $g(k) = \frac{1}{k^2}$ and note that

$$\int_X g \, \mathrm{d}\mu_\Sigma = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

by Example 2.4.2–**??**. Therefore, the hypotheses of the Dominated Convergence Theorem apply, and we have

$$\lim_{j\to\infty} \int_{\mathbb{Z}_{>0}} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma = \int_{\mathbb{Z}_{>0}} \lim_{j\to\infty} f_{\alpha_j} \, \mathrm{d}\mu_\Sigma = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} = \frac{\pi^2}{12},$$

where we look up the last sum. •

5.7.4 Integration with respect to probability measures

In Section 5.3.5 we introduced the notion of a probability space. In this section we investigate integration on probability spaces, giving a few results peculiar and useful for such measure spaces.

The following general result concerning how integrals behave under composition by certain classes of functions. Recall from Sections 3.1.6 and 3.2.6 the notion of a convex function. We shall use properties of convex functions we proved in those sections.

5.7.31 Theorem (Jensen's[11] inequality) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a finite measure space, let* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R})$*, and let* $\phi\colon \mathbb{R} \to \mathbb{R}$ *be convex. Then*

$$\phi\Big(\int_{\mathsf{X}} \mathsf{f}\,\mathrm{d}\mu\Big) \le \int_{\mathsf{X}} (\phi \circ \mathsf{f})\,\mathrm{d}\mu.$$

Proof From Proposition 3.2.30(**??**) we have

$$\phi'(y_0+)(y - y_0) + \phi(y_0) \le \phi(y)$$

for every $y \in \mathbb{R}$. Let $x \in X$ and let us take

$$y_0 = \int_X f\,\mathrm{d}\mu, \quad y = f(x),$$

so that the above inequality reads

$$\phi(y_0) \le \phi \circ f(x) - \phi'(y_0+)(f(x) - y_0).$$

By Proposition 5.7.19 we have

$$\int_X \phi(y_0)\,\mathrm{d}\mu \le \int_X \phi \circ f\,\mathrm{d}\mu - \int_X \phi'(y_0+)(f - y_0)\,\mathrm{d}\mu.$$

Since μ is a probability measure (i.e., $\int_X \mathrm{d}\mu = 1$) and since the integral is linear, we have

$$\int_X \phi(y_0)\,\mathrm{d}\mu = \phi(y_0)$$

and

$$\int_X \phi'(y_0+)(f - y_0)\,\mathrm{d}\mu = \phi'(y_0+)\int_X f\,\mathrm{d}\mu - \phi'(y_0+)y_0 = 0.$$

This immediately gives the result. ■

The following version of Jensen's inequality is often useful. Here we make use of the Lebesgue integral on $\mathbb{R}$ discussed in detail in Section 5.9.

[11] Johan Ludwig William Valdemar Jensen (1859–1925) was a Danish telephone company employee who did some mathematics in his spare time.

5.7.32 Corollary (Jensen's inequality for integration on intervals) *Let* $[a,b] \subseteq \mathbb{R}$ *be a compact interval, let* $f \in \mathsf{L}^{(1)}(([a,b],\mathscr{L}([a,b]),\lambda_{[a,b]});\mathbb{R})$*, and let* $\phi\colon \mathbb{R}\to\mathbb{R}$ *be convex. Then*

$$\phi\Big(\int_{[a,b]} f\,d\lambda_{[a,b]}\Big) \le \frac{1}{b-a}\int_{[a,b]} \phi\circ((b-a)f)\,d\lambda_{[a,b]}.$$

Proof We shall use Riemann integral notation in the proof, cf. Notation 5.9.13. By the change of variable theorem, Theorem 5.9.36,

$$\int_a^b f(x)\,dx = \int_0^1 (b-a)f(a+(b-a)s)\,ds.$$

By Jensen's inequality above,

$$\phi\Big(\int_a^b f(x)\,dx\Big) \le \int_0^1 \phi(b-a)f(a+(b-a)s))\,ds = \frac{1}{b-a}\int_a^b \phi((b-a)f(x))\,dx,$$

which is the result. ■

Now we give a few characterisations of how a function deviates from its mean. For this, the following simple definition is useful.

5.7.33 Definition (Mean of a function) Let $(X,\mathscr{A},\mu)$ be a measure space and let $f \in \mathsf{L}^{(1)}((X,\mathscr{A},\mu);\overline{\mathbb{R}})$. The ***mean*** of f is

$$\mathrm{mean}(f) = \int_X f\,d\mu. \qquad \bullet$$

With this notion, we have the following results.

5.7.34 Theorem (Markov's[12] inequality) *Let* $(X,\mathscr{A},\mu)$ *be a measure space and let* $f \in \mathsf{L}^{(1)}((X,\mathscr{A},\mu);\overline{\mathbb{R}}_{\ge 0})$*. Then, for any* $a \in \mathbb{R}_{>0}$ *it holds that*

$$\mu(\{x \in X \mid f(x) \ge a\}) \le \frac{1}{a}\mathrm{mean}(f).$$

Proof Let us abbreviate

$$M_a = \{x \in X \mid f(x) \ge a\}.$$

Then, for every $x \in X$.

$$a \le a\chi_{M_a}(x) \le f(x)\chi_{M_a}(x) \le f(x).$$

Therefore, by Proposition 5.7.19,

$$\int_X (a\chi_{M_a})\,d\mu \le \int_{M_a} f\,d\mu_{M_a} \le \int_X f\,d\mu.$$

Dividing by a gives $\mu(M_a) \le \frac{1}{a}\mathrm{mean}(f)$, as desired. ■

Very often Markov's inequality gives rather course estimates, and moreover only applies to nonnegative-valued functions. In this respect, the following results are sometimes useful.

[12] Andrei Andreyevich Markov (1856–1922) did mathematical research in analysis, and was one of the pioneers in the early development of what we now know as probability theory. He also involved himself in the political turmoil in which Russia was involved during his lifetime.

5.7.35 Theorem (General Chebychev[13] inequality) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space, let* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$*, and let* $\phi\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ *be such that* $\phi(\mathsf{y}_1) \leq \mathsf{g}(\mathsf{y}_2)$ *for all* $\mathsf{y}_1, \mathsf{y}_2 \in \mathrm{image}(\mathsf{f})$ *with* $\mathsf{y}_1 < \mathsf{y}_2$*. Then, for any* $\mathsf{a} \in \mathbb{R}$ *for which* $\phi(\mathsf{a}) \in \mathbb{R}$*, it holds that*

$$\mu(\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) \geq \mathsf{a}\}) \leq \frac{1}{\phi(\mathsf{a})}\mathrm{mean}(\phi \circ \mathsf{f}).$$

Proof Let

$$M_a = \{x \in X \mid f(x) \geq a\}.$$

Then, for any $x \in X$,

$$\phi(a)\chi_{M_a}(x) \leq \phi \circ f(x)\chi_{M_a}(x) \leq \phi \circ f(x),$$

noting that $\phi \circ f(x) \geq \phi(a)$ for $x \in M_a$ since ϕ is monotonically increasing. Using Proposition 5.7.19, just as in the proof of Markov's inequality, we have

$$\phi(a)\mu(M_a) \leq \mathrm{mean}(\phi \circ f),$$

as desired. ■

The usual form of Chebychev's inequality is the following.

5.7.36 Corollary (Usual form of Chebychev's inequality) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}_{\geq 0})$*. Then, for any* $\mathsf{a} \in \mathbb{R}_{>0}$ *it holds that*

$$\mu(\{\mathsf{x} \in \mathsf{X} \mid |\mathsf{f}(\mathsf{x})| \geq \mathsf{a}\}) \leq \frac{1}{\mathsf{a}^2}\int_{\mathsf{X}} \mathsf{f}^2\,\mathrm{d}\mu.$$

Proof Applying the general form of Chebychev's inequality with

$$\phi(y) = \begin{cases} y^2, & y \in \overline{\mathbb{R}}_{\geq 0}, \\ 0, & \text{otherwise} \end{cases}$$

and replacing f with $|f|$ gives the result. ■

Our final result of this form is the following result which follows from our general for the Chebychev inequality. For $c \in \mathbb{R}$ let us denote $\exp_c\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}_{\geq 0}$ by

$$\exp_c(y) = \begin{cases} \mathrm{e}^{cy}, & y \in \mathbb{R}_{\geq 0}, \\ \lim_{y \to -\infty} \mathrm{e}^{cy}, & y = -\infty, \\ \lim_{y \to \infty} \mathrm{e}^{cy}, & y = \infty, \end{cases}$$

allowing that one of the limits will be ∞. With this notation we have the following result.

[13] Pafnuty Lvovich Chebyshev (1821–1894) was a Russian mathematician, making contributions to the areas of analysis, number theory, and approximation theory, and was one of the early researchers in the area of modern probability theory.

5.7.37 Corollary (Chernoff's[14] inequality) *Let* $(X, \mathscr{A}, \mu)$ *be a measure space and let* $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$*. Then, for any* $a, c \in \mathbb{R}_{>0}$*, it holds that*

$$\mu(\{x \in X \mid f(x) \geq a\}) \leq e^{-ca} \int_X \exp_c \circ f \, d\mu.$$

Proof Applying the general form of Chebychev's inequality with $\phi = \exp_c$ gives the result. ■

Note that it might very well be the case that the right-hand side of either of the inequalities in the preceding two corollaries will be infinite. In this case the inequalities hold vacuously, and so do not give useful information.

5.7.5 Topological characterisations of limit theorems[15]

It turns out that there is a very simple way to restate usual version of the Dominated Convergence Theorem using the notion of a limit structure for almost everywhere pointwise convergence from Theorem 5.6.51. For this purpose, it is advantageous to have at hand two versions of the Dominated Convergence Theorem. One is that stated as Theorem 5.7.28, and the other, an "everywhere" rather than an "almost everywhere" version, being the following.

5.7.38 Theorem ("Everywhere" Dominated Convergence Theorem) *Let* $(X, \mathscr{A}, \mu)$ *be a measure space and let* $(f_j)_{j \in \mathbb{Z}_{>0}}$ *be a sequence in* $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ *having the following properties:*

(i) the limit $f(x) = \lim_{j\to\infty} f_j(x)$ *exists for every* $x \in X$*;*

(ii) there exists $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}}_{\geq 0})$ *such that, for every* $x \in X$, $|f_j(x)| \leq g(x)$ *for every* $j \in \mathbb{Z}_{>0}$*.*

Then the functions f *and* f_j*,* $j \in \mathbb{Z}_{>0}$*, are integrable and*

$$\int_X f \, d\mu = \lim_{j\to\infty} \int_X f_j \, d\mu.$$

Proof This follows immediately from Theorem 5.7.28. ■

Our objective is to restate the "everywhere" and "almost everywhere" versions of the Dominated Convergence Theorem in topological terms. First let us consider the "everywhere" version of the Dominated Convergence Theorem, Theorem 5.7.38. In this case we use the topology $\mathscr{C}_p$ of pointwise convergence on $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ described in Section **??**. Note that Proposition 5.6.18 implies that $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ is a sequentially closed subspace of $\mathbb{R}^X$ using this topology. Let us say that a subset $A \subseteq \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ is $\mathbf{C_p}$***-sequentially closed*** if every $\mathscr{C}_p$ convergent sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ in A converges to a function in A. A subset $B \subseteq \mathbb{R}^X$ is $\mathbf{C_p}$***-bounded*** if, for every sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ in B and every sequence $(a_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathbb{R}$ converging to

[14] Herman Chernoff, born in New York in 1923, is an American statistician.

[15] The results in this section are not used in an essential way elsewhere in the text, except in Section 5.9.11.

0, the sequence $(a_j f_j)_{j\in\mathbb{Z}_{>0}}$ converges to the zero function in the C_p-topology. This notion of boundedness may look strange at present. We shall examine the general context from which this definition is derived in ***missing stuff***.

The following result characterises C_p-bounded sets.

5.7.39 Proposition (Characterisation of C_p-bounded functions) *Let* X *be a set. A subset* $\mathrm{B} \subseteq \mathbb{R}^{\mathrm{X}}$ *is* C_p*-bounded if and only if there exists a nonnegative-valued* $\mathrm{g} \in \mathbb{R}^{\mathrm{X}}$ *such that*

$$\mathrm{B} \subseteq \{\mathrm{f} \in \mathbb{R}^{\mathrm{X}} \mid |\mathrm{f}(\mathrm{x})| \le \mathrm{g}(\mathrm{x}) \textit{ for every } \mathrm{x} \in \mathrm{X}\}.$$

Proof Suppose that there exists a nonnegative-valued $g \in \mathbb{R}^X$ such that $|f(x)| \le g(x)$ for every $x \in X$ if $f \in B$. Let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in B and let $(a_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$ converging to 0. If $x \in X$ then

$$\lim_{j\to\infty} |a_j f_j(x)| \le \lim_{j\to\infty} |a_j| g(x) = 0,$$

which gives C_p-convergence of the sequence $(a_j f_j)_{j\in\mathbb{Z}_{>0}}$ to zero.

Next suppose that there exists no nonnegative-valued function $g \in \mathbb{R}^X$ such that $|f(x)| \le g(x)$ for every $x \in X$ if $f \in B$. This means that there exists $x_0 \in X$ such that, for every $M \in \mathbb{R}_{>0}$, there exists $f \in B$ such that $|f(x_0)| > M$. Let $(a_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$ converging to 0 and such that $a_j \neq 0$ for every $j \in \mathbb{Z}_{>0}$. Then let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in B such that $|f_j(x_0)| > |a_j^{-1}|$ for every $j \in \mathbb{Z}_{>0}$. Then $|a_j f_j(x_0)| > 1$ for every $j \in \mathbb{Z}_{>0}$, implying that the sequence $(a_j f_j)_{j\in\mathbb{Z}_{>0}}$ cannot C_p-converge to zero. Thus B is not C_p-bounded. ■

With the preceding development, we can now state the "everywhere" Dominated Convergence Theorem in terms of the C_p-topology.

5.7.40 Theorem (Topological "everywhere" Dominated Convergence Theorem) *If* $(\mathrm{X}, \mathscr{A}, \mu)$ *is a measure space then* C_p*-bounded subsets of* $\mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \mathbb{R})$ *are* C_p*-sequentially closed.*

Proof This follows immediately from Theorem 5.7.38 and the definitions of the terms involved. ■

Now we turn to the "almost everywhere" Dominated Convergence Theorem. Here matters are possibly (and often) complicated by the fact that almost everywhere pointwise convergence is not topological, as shown in Proposition 5.6.48. However, we can effectively replace the rôle of the C_p-topology above with the $\mathscr{L}_\mu$-limit structure. To this end, let us say that a subset $A \subseteq \mathsf{L}^0((X, \mathscr{A}); \mathbb{R})$ is ***$\mathscr{L}_\mu$-sequentially closed*** if every $\mathscr{L}_\mu$ convergent sequence $([f_j])_{j\in\mathbb{Z}_{>0}}$ in A converges to an equivalence class of functions in A. A subset $B \subseteq \mathsf{L}^0((X, \mathscr{A}); \mathbb{R})$ is ***$\mathscr{L}_\mu$-bounded*** if, for every sequence $([f_j])_{j\in\mathbb{Z}_{>0}}$ in B and every sequence $(a_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ converging to 0, the sequence $([a_j f_j])_{j\in\mathbb{Z}_{>0}}$ converges to the zero equivalence class in the $\mathscr{L}_\mu$-topology.

The following result characterises $\mathscr{L}_\mu$-bounded sets.

5.7.41 Proposition *A subset* $\mathrm{B} \subseteq \mathsf{L}^0((\mathrm{X}, \mathscr{A}); \mathbb{R})$ *is* $\mathscr{L}_\mu$*-bounded if and only if there exists a nonnegative-valued* $\mathrm{g} \in \mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \mathbb{R})$ *such that*

$$\mathrm{B} \subseteq \{[\mathrm{f}] \in \mathsf{L}^0((\mathrm{X}, \mathscr{A}); \mathbb{R}) \mid |\mathrm{f}(\mathrm{x})| \le \mathrm{g}(\mathrm{x}) \textit{ for almost every } \mathrm{x} \in \mathrm{X}\}.$$

Proof We first observe that the condition that $|f(x)| \leq g(x)$ for almost every $x \in X$ is independent of the choice of representative f from the equivalence class $[f]$.

Suppose that there exists a nonnegative-valued $g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ such that, if $[f] \in B$, then $|f(x)| \leq g(x)$ for almost every $x \in X$. Let $([f_j])_{j \in \mathbb{Z}_{>0}}$ be a sequence in B and let $(a_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$ converging to zero. For $j \in \mathbb{Z}_{>0}$ define

$$A_j = \{x \in X \mid |f_j(x)| \leq g(x)\}.$$

Note that if $x \in X \setminus (\cup_{j \in \mathbb{Z}_{>0}} A_j)$ then

$$\lim_{j \to \infty} |a_j f_j(x)| \leq \lim_{j \to \infty} |a_j| g(x) = 0.$$

Since $\mu(\cup_{j \in \mathbb{Z}_{>0}} A_j) = 0$ this implies that the sequence $(a_j[f_j])_{j \in \mathbb{Z}_{>0}}$ is $\mathscr{L}_\mu$-convergent to zero. One may show that this argument is independent of the choice of representatives f_j from the equivalence classes $[f_j]$, $j \in \mathbb{Z}_{>0}$.

Conversely, suppose that there exists no nonnegative-valued function $g \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ such that, for every $[f] \in B$, $|f(x)| \leq g(x)$ for almost every $x \in X$. This means that there exists a set $E \subseteq X$ of positive measure such that, for any $M \in \mathbb{R}_{>0}$, there exists $[f] \in B$ such that $|f(x)| > M$ for almost every $x \in E$. Let $(a_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$ converging to 0 and such that $a_j \neq 0$ for every $j \in \mathbb{Z}_{>0}$. Then let $([f_j])_{j \in \mathbb{Z}_{>0}}$ be a sequence in B such that $|f_j(x)| > |a_j^{-1}|$ for almost every $x \in E$ and for every $j \in \mathbb{Z}_{>0}$. Define

$$A_j = \{x \in E \mid |f_j(x)| > |a_j^{-1}|\}.$$

If $x \in E \setminus (\cup_{j \in \mathbb{Z}_{>0}} A_j)$ then $|a_j f_j(x)| > 1$ for every $j \in \mathbb{Z}_{>0}$. Since $\mu(E \setminus (\cup_{j \in \mathbb{Z}_{>0}} A_j)) > 0$ it follows that $(a_j[f_j])_{j \in \mathbb{Z}_{>0}}$ cannot $\mathscr{L}_\mu$-converge to zero, and so B is not $\mathscr{L}_\mu$-bounded. ■

We can then state the following characterisation of the "almost everywhere" Dominated Convergence Theorem. We denote by $\mathsf{L}^1((X, \mathscr{A}, \mu); \mathbb{R})$ the image of $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R})$ under the projection from $\mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R})$ to $\mathsf{L}^0((X, \mathscr{A}); \mathbb{R})$. Thus elements of $\mathsf{L}^1((X, \mathscr{A}, \mu); \mathbb{R})$ are equivalence classes of integrable $\mathbb{R}$-valued functions under the equivalence relation of almost everywhere equality. The space $\mathsf{L}^1((X, \mathscr{A}, \mu); \mathbb{R})$ will be studied in detail as part of Section 6.7.8.

5.7.42 Theorem (Limit structure "almost everywhere" Dominated Convergence Theorem) *If* $(\mathrm{X}, \mathscr{A}, \mu)$ *is a measure space then* $\mathscr{L}_\mu$*-bounded subsets of* $\mathsf{L}^1((\mathrm{X}, \mathscr{A}, \mu); \mathbb{R})$ *are* $\mathscr{L}_\mu$*-sequentially closed.*

Proof This follows immediately from Theorem 5.7.28 and the definitions of the terms involved. ■

5.7.6 Image measure and integration by image measure

In this section we provide the definition of a measure induced by a map. We shall not use this construction frequently, but it does arise, for example, in parts of our discussion of convolution in Chapter 11.***missing stuff***

The construction is as follows.

5.7.43 Proposition (Characterisation of image measure) *Let $(\mathsf{X}, \mathscr{A}, \mu)$ be a measure space, let $(\mathsf{Y}, \mathscr{B})$ be a measurable space, and let $\phi \colon \mathsf{X} \to \mathsf{Y}$ be a $(\mathscr{A}, \mathscr{B})$-measurable map. If we define $\mu\phi^{-1} \colon \mathscr{B} \to \overline{\mathbb{R}}_{\geq 0}$ by $\mu\phi^{-1}(\mathsf{B}) = \mu(\phi^{-1}(\mathsf{B}))$, then $(\mathsf{Y}, \mathscr{B}, \mu\phi^{-1})$ is a measure space.*

Proof Since $\phi^{-1}(\emptyset) = \emptyset$ we have $\mu\phi^{-1}(\emptyset) = 0$. Now let $(B_j)_{j \in \mathbb{Z}_{>0}}$ be a pairwise disjoint family of subsets from $\mathscr{B}$. We claim that $(\phi^{-1}(B_j))_{j \in \mathbb{Z}_{>0}}$ is pairwise disjoint. This follows since $\phi^{-1}(B_j) \cap \phi^{-1}(B_k) = \phi^{-1}(B_j \cap B_k)$ by Proposition 1.3.5. It, therefore, follows that

$$\sum_{j=1}^{\infty} \mu\phi^{-1}(B_j) = \sum_{j=1}^{\infty} \mu(\phi^{-1}(B_j)) = \mu\Big(\bigcup_{j \in \mathbb{Z}_{>0}} \phi^{-1}(B_j)\Big) = \mu\phi^{-1}\Big(\bigcup_{j \in \mathbb{Z}_{>0}} B_j\Big),$$

again with an application of Proposition 1.3.5. ■

The measure $\mu\phi^{-1}$ has a name.

5.7.44 Definition (Image measure) For $(X, \mathscr{A}, \mu)$, $(Y, \mathscr{B})$, and ϕ as in Proposition 5.7.43, the measure $\mu\phi^{-1}$ is the ***image measure*** of μ by ϕ. •

One can characterise the functions integrable by the image measure.

5.7.45 Proposition (Integration by the image measure) *Let $(\mathsf{X}, \mathscr{A}, \mu)$ be a measure space, let $\phi \colon \mathsf{X} \to \mathsf{Y}$ be a $(\mathsf{Y}, \mathscr{B})$ be a $(\mathscr{A}, \mathscr{B})$-measurable map, and let $\mu\phi^{-1}$ be the image measure of μ by ϕ. Then $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{Y}, \mathscr{B}); \overline{\mathbb{R}})$ is integrable with respect to $\mu\phi^{-1}$ if and only if $\mathsf{f} \circ \phi$ is integrable with respect to μ. Moreover, if $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \mu\phi^{-1}); \overline{\mathbb{R}})$ then we have*

$$\int_{\mathsf{Y}} \mathsf{f} \, \mathrm{d}(\mu\phi^{-1}) = \int_{\mathsf{X}} (\mathsf{f} \circ \phi) \, \mathrm{d}\mu.$$

Proof Suppose that f is $\mu\phi^{-1}$-integrable. By Proposition 5.6.6 this means that f is $(\mathscr{B}, \mathscr{B}(\overline{\mathbb{R}}))$-measurable. Since ϕ is $(\mathscr{A}, \mathscr{B})$-measurable, it follows easily that $f \circ \phi$ is $(\mathscr{A}, \mathscr{B}(\overline{\mathbb{R}}))$-measurable, and so measurable.

Now let $B \in \mathscr{B}$ and note that $\chi_B \circ \phi = \chi_{\phi^{-1}(B)}$, as can be directly verified. Therefore,

$$\int_Y \chi_B \, \mathrm{d}(\mu\phi^{-1}) = \mu\phi^{-1}(B) = \mu(\phi^{-1}(B)) = \int_X \chi_{\phi^{-1}(B)} \, \mathrm{d}\mu = \int_X \chi_B \circ \phi \, \mathrm{d}\mu.$$

By linearity of the integral, Proposition 5.7.17, this implies that if $f \in \mathsf{L}^{(0)}((Y, \mathscr{B}); \overline{\mathbb{R}})$ is a simple function we have

$$\int_Y f \, \mathrm{d}(\mu\phi^{-1}) = \int_X (f \circ \phi) \, \mathrm{d}\mu. \tag{5.18}$$

If $f \in \mathsf{L}^{(1)}((Y, \mathscr{B}, \mu\phi^{-1}); \overline{\mathbb{R}}_{\geq 0})$ then by Proposition 5.6.39 there exists a sequence of monotonically increasing simple functions $(g_j)_{j \in \mathbb{Z}_{>0}}$ such that $f(y) = \lim_{j \to \infty} g_j(y)$ for each $y \in Y$. The sequence $(g_j \circ \phi)_{j \in \mathbb{Z}_{>0}}$ is then itself a sequence of monotonically increasing functions such that $f \circ \phi(x) = \lim_{j \to \infty} g_j \circ \phi(x)$. By the Monotone Convergence Theorem, (5.18) then holds for $f \in \mathsf{L}^{(1)}((Y, \mathscr{B}, \mu\phi^{-1}); \overline{\mathbb{R}}_{\geq 0})$. For general integrable functions, breaking the function f into its positive and negative parts and using linearity of the integral gives (5.18) in this case. This shows that if $f \in \mathsf{L}^{(1)}((Y, \mathscr{B}, \mu\phi^{-1}); \overline{\mathbb{R}})$ then $f \circ \phi \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ and the functions have equal integrals.

The argument above also clearly shows that if $f \circ \phi \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ then $f \in \mathsf{L}^{(1)}((Y, \mathscr{B}, \mu\phi^{-1}); \overline{\mathbb{R}})$, as desired. ■

5.7.7 The integral for $\mathbb{C}$- and vector-valued functions

Thus far, we have always assumed that functions take values in $\overline{\mathbb{R}}$ or subsets of $\overline{\mathbb{R}}$. Some of the time, however, we wish to integrate functions that are vector-valued, or particularly $\mathbb{C}$-valued. The extension to these sorts of functions is easily made, and in this section we write the (hopefully) expected results. The reader will wish to recall our discussion in Section 5.6.4 of measurable vector-valued functions.

We begin with the definitions.

5.7.46 Definition (Integrable vector-valued function) For a measure space $(X, \mathscr{A}, \mu)$, a function $\boldsymbol{f}\colon X \to \mathbb{R}^n$ is ***integrable*** if its components $f_1, \dots, f_n$ are integrable in the sense of Definition 5.7.8. The ***integral*** of an integrable function $\boldsymbol{f}\colon X \to \mathbb{R}^n$ is

$$\int_X \boldsymbol{f}\,\mathrm{d}\mu = \Big(\int_X f_1\,\mathrm{d}\mu, \dots, \int_X f_n\,\mathrm{d}\mu\Big).$$

We denote the set of integrable $\mathbb{R}^n$-valued maps by $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$. •

The following result gives a useful characterisation of the integrability of $\mathbb{R}^n$-valued functions.

5.7.47 Proposition (Characterisation of vector-valued integrable functions) *For a measure space* $(\mathrm{X}, \mathscr{A}, \mu)$ *and* $\mathbf{f} \in \mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \mathbb{R}^{\mathrm{n}})$*, the following statements are equivalent:*

(i) $\mathbf{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathrm{n}})$*;*

(ii) the $\mathbb{R}$*-valued function* $\mathrm{x} \mapsto \|\mathbf{f}\|_{\mathbb{R}^{\mathrm{n}}}(\mathrm{x})$ *is integrable.*

Moreover, if either of the above equivalent conditions holds, then

$$\Big\|\int_{\mathrm{X}} \mathbf{f}\,\mathrm{d}\mu\Big\|_{\mathbb{R}^{\mathrm{n}}} \le \int_{\mathrm{X}} \|\mathbf{f}\|_{\mathbb{R}^{\mathrm{n}}}\,\mathrm{d}\mu.$$

Proof (i) $\implies$ (ii) By Proposition 5.6.11 and Corollary 5.6.33 it follows that $x \mapsto \|\boldsymbol{f}(x)\|_{\mathbb{R}^n}$ is measurable. From Lemma **??** we have

$$\|\boldsymbol{f}(x)\|_{\mathbb{R}^n} \le |f_1(x)| + \dots + |f_n(x)|$$

for every $x \in X$. Therefore, by Propositions 5.7.17 and 5.7.19,

$$\int_X \|\boldsymbol{f}\|_{\mathbb{R}^n}\,\mathrm{d}\mu \le \int_X |f_1|\,\mathrm{d}\mu + \dots + \int_X |f_n|\,\mathrm{d}\mu < \infty,$$

giving the result.

(ii) $\implies$ (i) From Lemma **??** we have

$$|f_1(x)| + \dots + |f_n(x)| \le \sqrt{n}\|\boldsymbol{f}(x)\|_{\mathbb{R}^n}$$

for every $x \in X$. Therefore, by Proposition 5.7.19, for each $j \in \{1, \dots, n\}$ we have

$$\int_X |f_j|\,\mathrm{d}\mu \le \int_X \|\boldsymbol{f}\|_{\mathbb{R}^n}\,\mathrm{d}\mu < \infty,$$

as desired.

Now we prove the final assertion of the proposition. The inequality obviously holds if $\int_X \boldsymbol{f}\,\mathrm{d}\mu = \boldsymbol{0}$, so we may suppose that $\int_X \boldsymbol{f}\,\mathrm{d}\mu \neq \boldsymbol{0}$. Let $\boldsymbol{u} \in \mathbb{R}^n$ be such that $\|\boldsymbol{u}\|_{\mathbb{R}^n} = 1$ and

$$\int_X \boldsymbol{f}\,\mathrm{d}\mu = \boldsymbol{u}\Big\|\int_X \boldsymbol{f}\,\mathrm{d}\mu\Big\|_{\mathbb{R}^n}.$$

Therefore, using linearity of the integral and the fact that $\langle \boldsymbol{u}, \boldsymbol{u}\rangle_{\mathbb{R}^n} = 1$,

$$\int_X \langle \boldsymbol{u}, \boldsymbol{f}\rangle_{\mathbb{R}^n}\,\mathrm{d}\mu = \Big\langle \boldsymbol{u}, \int_X \boldsymbol{f}\,\mathrm{d}\mu\Big\rangle_{\mathbb{R}^n} = \Big\|\int_X \boldsymbol{f}\,\mathrm{d}\mu\Big\|_{\mathbb{R}^n}.$$

Since $|u_j| \leq 1$ for each $j \in \{1, \dots, n\}$ we can use the Cauchy–Bunyakovsky–Schwarz inequality and Lemma **??** to get

$$\langle \boldsymbol{u}, \boldsymbol{f}(x)\rangle_{\mathbb{R}^n} \leq |\langle \boldsymbol{u}, \boldsymbol{f}(x)\rangle_{\mathbb{R}^n}| \leq \|\boldsymbol{u}\|_{\mathbb{R}^n}\|\boldsymbol{f}(x)\|_{\mathbb{R}^n} = \|\boldsymbol{f}(x)\|_{\mathbb{R}^n}.$$

Therefore, by Proposition 5.7.19,

$$\Big\|\int_X \boldsymbol{f}\,\mathrm{d}\mu\Big\|_{\mathbb{R}^n} \leq \int_X \|\boldsymbol{f}(x)\|_{\mathbb{R}^n}\,\mathrm{d}\mu,$$

as desired. ■

This result has the following immediate and useful corollary which gives an easy means of checking the integrability of a vector-valued function.

5.7.48 Corollary (Integrability of vector-valued functions) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *be a measure space and let* $\mathbf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathrm{n}})$ *and* $\mathrm{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}), \mathbb{R}_{\geq 0})$ *satisfy* $\|\mathbf{f}(\mathrm{x})\|_{\mathbb{R}^{\mathrm{n}}} \leq \mathrm{g}(\mathrm{x})$ *for almost every* $\mathrm{x} \in \mathsf{X}$*. Then* $\mathbf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathrm{n}})$ *if* $\mathrm{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}_{\geq 0})$ *and, in this case,*

$$\Big\|\int_{\mathsf{X}} \mathbf{f}\,\mathrm{d}\mu\Big\|_{\mathbb{R}^{\mathrm{n}}} \leq \int_{\mathsf{X}} \mathrm{g}\,\mathrm{d}\mu.$$

Of course, the preceding definition and characterisation of integrable vector-valued functions applies immediately to $\mathbb{C}$-valued functions, using the fact that $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic as $\mathbb{R}$-vector spaces.

5.7.49 Definition (Integrable $\mathbb{C}$-valued function) For a measure space $(X, \mathscr{A}, \mu)$, a function $f\colon X \to \mathbb{C}$ is ***integrable*** if the $\mathbb{R}$-valued functions

$$\mathrm{Re}(f)\colon x \mapsto \mathrm{Re}(f(x)), \quad \mathrm{Im}(f)\colon x \mapsto \mathrm{Im}(f(x))$$

are integrable in the sense of Definition 5.7.8. The ***integral*** of an integrable function $f\colon X \to \mathbb{C}$ is

$$\int_X f\,\mathrm{d}\mu = \Big(\int_X \mathrm{Re}(f)\,\mathrm{d}\mu, \int_X \mathrm{Im}(f)\,\mathrm{d}\mu\Big).$$

We denote the set of integrable $\mathbb{C}$-valued maps by $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{C})$. •

Following immediately from Proposition 5.7.47 is the following result.

5.7.50 Corollary (Characterisation of $\mathbb{C}$-valued integrable functions) *For a measure space* $(X, \mathscr{A}, \mu)$ *and* $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{C})$*, the following statements are equivalent:*

(i) $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{C})$*;*

(ii) the $\mathbb{R}$*-valued function* $x \mapsto |f|(x)$ *is integrable.*

Moreover, if either of the above equivalent statements holds then

$$\left|\int_X f \, d\mu\right| \le \int_X |f| \, d\mu.$$

Most of the properties of the integral generalise to vector- or $\mathbb{C}$-valued integrals. For completeness we record the results explicitly for $\mathbb{R}^n$-valued functions, noting that these results apply immediately to $\mathbb{C}$-valued functions.

The following result is fundamental and often used without explicit mention.

5.7.51 Proposition (Integrals of functions agreeing almost everywhere) *Let* $(X, \mathscr{A}, \mu)$ *be a measure space and let* $\mathbf{f}, \mathbf{g} \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R}^n)$ *have the property that* $\mathbf{f}(x) = \mathbf{g}(x)$ *for almost every* $x \in X$*. Then the integral of* $\mathbf{f}$ *exists if and only if the integral of* $\mathbf{g}$ *exists, and if either integral exists then we have*

$$\int_X \mathbf{f} \, d\mu = \int_X \mathbf{g} \, d\mu.$$

Proof This follows immediately from Proposition 5.7.11, along with the definition of the integral for vector-valued functions. ■

Next let us see that the vector-valued integral is linear.

5.7.52 Proposition (Algebraic operations on integrable functions) *For a measure space* $(X, \mathscr{A}, \mu)$*, for* $\mathbf{f}, \mathbf{g} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$*, and for* $\alpha \in \mathbb{R}$*, the following statements hold:*

(i) $\mathbf{f} + \mathbf{g} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$ *and*

$$\int_X (\mathbf{f} + \mathbf{g}) \, d\mu = \int_X \mathbf{f} \, d\mu + \int_X \mathbf{g} \, d\mu;$$

(ii) $\alpha \mathbf{f} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$ *and*

$$\int_X \alpha \mathbf{f} \, d\mu = \alpha \int_X \mathbf{f} \, d\mu.$$

Proof This follows directly from Proposition 5.7.17 and the definition of the vector-valued integral. ■

It is also useful to know that the integral of $\mathbb{C}$-valued functions is $\mathbb{C}$-linear.

5.7.53 Corollary (Linearity of the $\mathbb{C}$ integral) *For a measure space* $(X, \mathscr{A}, \mu)$*, for* $f, g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{C})$*, and for* $\alpha \in \mathbb{C}$*, the following statements hold:*

(i) $f + g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{C})$ *and*

$$\int_X (f + g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu;$$

(ii) $\alpha f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{C})$ *and*

$$\int_X \alpha f\, d\mu = \alpha \int_X f\, d\mu.$$

Proof The first assertion is a special case of the first assertion of Proposition 5.7.52. The second assertion also follows from Proposition 5.7.52 since

$$\operatorname{Re}(\alpha f) = \operatorname{Re}(\alpha)\operatorname{Re}(f) - \operatorname{Im}(\alpha)\operatorname{Im}(f), \qquad \operatorname{Im}(\alpha f) = \operatorname{Re}(\alpha)\operatorname{Im}(f) + \operatorname{Im}(\alpha)\operatorname{Re}(f). \qquad \blacksquare$$

For integrating vector-valued functions over disjoint subsets, we have the following result.

5.7.54 Proposition (Breaking the integral in two) *Let* $(X, \mathscr{A}, \mu)$*, let* $A, B \in \mathscr{A}$ *be sets such that* $X = A \overset{\circ}{\cup} B$*, and let* $\mathbf{f} \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R}^n)$*. Then* $\mathbf{f} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$ *if and only if* $\mathbf{f}|A \in \mathsf{L}^{(1)}((A, \mathscr{A}_A, \mu|\mathscr{A}_A); \mathbb{R}^n)$ *and* $\mathbf{f}|B \in \mathsf{L}^{(1)}((B, \mathscr{A}_B, \mu|\mathscr{A}_B); \mathbb{R}^n)$*. Furthermore, if* $\mathbf{f} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$ *then we have*

$$\int_X \mathbf{f}\, d\mu = \int_A (\mathbf{f}|A)\, d\mu_A + \int_B (\mathbf{f}|B)\, d\mu_B.$$

Proof Thus follows from Proposition 5.7.22, along with the definition of the integral for vector-valued functions. ■

As in the scalar case, this result has the following corollary.

5.7.55 Corollary (Breaking the integral almost in two) *Let* $(X, \mathscr{A}, \mu)$ *be a complete measure space, let* $A, B \in \mathscr{A}$ *be such that* $\mu(A \cap B) = 0$ *and such that* $X = A \cup B$*, and let* $\mathbf{f} \in \mathsf{L}^{(0)}((X, \mathscr{A}); \mathbb{R}^n)$*. Then* $\mathbf{f} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$ *if and only if* $\mathbf{f}|A \in \mathsf{L}^{(1)}((A, \mathscr{A}_A, \mu|\mathscr{A}_A); \mathbb{R}^n)$ *and* $\mathbf{f}|B \in \mathsf{L}^{(1)}((B, \mathscr{A}_B, \mu|\mathscr{A}_B); \mathbb{R}^n)$*. Furthermore, if* $\mathbf{f} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$ *then we have*

$$\int_X \mathbf{f}\, d\mu = \int_A (\mathbf{f}|A)\, d\mu_A + \int_B (\mathbf{f}|B)\, d\mu_B.$$

Proof This follows from Proposition 5.7.23, along with the definition of the vector-valued integral. ■

Finally, we can also state a version of the Dominated Convergence Theorem for vector-valued integrals.

5.7.56 Theorem (Vector-valued Dominated Convergence Theorem) *Let $(\mathsf{X}, \mathscr{A}, \mu)$ be a measure space and let $(\mathbf{f}_{\mathsf{j}})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}})$ having the following properties:*

(i) the limit $\mathbf{f}(\mathsf{x}) = \lim_{\mathsf{j}\to\infty} \mathbf{f}_{\mathsf{j}}(\mathsf{x})$ exists for almost every $\mathsf{x} \in \mathsf{X}$;

(ii) there exists $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}_{\geq 0})$ such that, for almost every $\mathsf{x} \in \mathsf{X}$, $\|\mathbf{f}_{\mathsf{j}}(\mathsf{x})\|_{\mathbb{R}^{\mathsf{n}}} \leq \mathsf{g}(\mathsf{x})$ for every $\mathsf{j} \in \mathbb{Z}_{>0}$.

Then the functions $\mathbf{f}$ and $\mathbf{f}_{\mathsf{j}}$, $\mathsf{j} \in \mathbb{Z}_{>0}$, are integrable and

$$\int_{\mathsf{X}} \mathbf{f}\,\mathrm{d}\mu = \lim_{\mathsf{j}\to\infty} \int_{\mathsf{X}} \mathbf{f}_{\mathsf{j}}\,\mathrm{d}\mu.$$

Proof For $k \in \{1, \dots, n\}$ denote by f_k the kth component of $\boldsymbol{f}$ and by $f_{j,k}$ the kth component of $\boldsymbol{f}_j$, $j \in \mathbb{Z}_{>0}$. Then, for almost every $x \in X$, we have

$$|f_k(x)| \leq \|\boldsymbol{f}(x)\|_{\mathbb{R}^n} \leq g(x), \qquad k \in \{1, \dots, n\},$$
$$|f_{j,k}(x)| \leq \|\boldsymbol{f}_j(x)\|_{\mathbb{R}^n} \leq g(x), \qquad k \in \{1, \dots, n\},\ j \in \mathbb{Z}_{>0}.$$

This gives integrability of $\boldsymbol{f}$ and $\boldsymbol{f}_j$, $j \in \mathbb{Z}_{>0}$, by definition of the vector-valued integral. The final equality of the theorem now follows from the scalar Dominated Convergence Theorem, Theorem 5.7.28. ■

5.7.8 Integration with respect to signed, complex, and vector measures

In this section to this point we have talked solely about positive measure spaces. Let us now see how signed, complex, and vector measure spaces arise in the integration story.

We begin by indicating how one can define integrals with respect to signed, complex, and vector measures. Here we use the Jordan decomposition of such measures in an essential way. Let us consider first the case where $(X, \mathscr{A}, \mu)$ is a signed measure space.

5.7.57 Definition (Integration with respect to a signed measure) For a signed measure space $(X, \mathscr{A}, \mu)$ let $\mu = \mu_+ - \mu_-$ be the Jordan decomposition of μ. For $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$, we have the following definitions.

(i) If neither of the conditions

(a) $\int_X f\,\mathrm{d}\mu_+ = \infty$ and $\int_X f\,\mathrm{d}\mu_- = \infty$ and

(b) $\int_X f\,\mathrm{d}\mu_+ = -\infty$ and $\int_X f\,\mathrm{d}\mu_- = -\infty$

holds, then the integral of f with respect to μ ***exists*** and is given by

$$\int_X f\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu_+ - \int_X f\,\mathrm{d}\mu_-,$$

this being the ***integral*** of f with respect to μ.

(ii) If either of the two conditions from part (i) hold then the integral of f with respect to μ ***does not exist***.

(iii) If $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_+); \overline{\mathbb{R}})$ and $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_-); \overline{\mathbb{R}})$ then f is ***integrable*** with respect to μ.

For a subset $I \subseteq \overline{\mathbb{R}}$ we denote the set of I-valued functions integrable with respect to μ by $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); I)$, or simply by $\mathsf{L}^{(1)}(X; I)$ if $\mathscr{A}$ and μ are understood. •

Using this definition of integrability and integral for signed measures, it is straightforward to define the corresponding notions for complex and vector measures. The essential idea is that a complex measure μ can be written as

$$\mu = \mathrm{Re}(\mu) + \mathrm{i}\,\mathrm{Im}(\mu)$$

for finite signed measures $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$. For a vector measure $\boldsymbol{\mu}$ taking values in $\mathbb{R}^n$, we can write

$$\boldsymbol{\mu} = \mu_1 \boldsymbol{e}_1 + \cdots + \mu_n \boldsymbol{e}_n$$

for finite signed measures $\mu_1, \dots, \mu_j$ and where $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is the standard basis for $\mathbb{R}^n$.

5.7.58 Definition (Integration with respect to complex and vector measures) For a measurable space $(X, \mathscr{A})$ and for a complex measure μ on $\mathscr{A}$ and a vector measure $\boldsymbol{\mu}$ taking values in $\mathbb{R}^n$, write them as above. For $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$, we have the following definitions.

(i) the integral of f with respect to μ ***exists*** if the integrals of f with respect to $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$ exist, and is given by

$$\int_X f \,\mathrm{d}\mu = \Big(\int_X f \,\mathrm{d}(\mathrm{Re}(\mu))\Big) + \mathrm{i}\Big(\int_X f \,\mathrm{d}(\mathrm{Im}(\mu))\Big),$$

this being the ***integral*** of f with respect to μ.

(ii) the integral of f with respect to $\boldsymbol{\mu}$ ***exists*** if the integrals of f with respect to $\mu_1, \dots, \mu_n$ exist, and is given by

$$\int_X f \,\mathrm{d}\boldsymbol{\mu} = \Big(\int_X f \,\mathrm{d}\mu_1, \dots, \int_X f \,\mathrm{d}\mu_n\Big),$$

this being the ***integral*** of f with respect to $\boldsymbol{\mu}$.

(iii) If the integral of f does not exist with respect to at least one of $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$, then the integral of f ***does not exist***.

(iv) If the integral of f does not exist with respect to at least one of $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$, then the integral of f ***does not exist***.

(v) If $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mathrm{Re}(\mu)); \overline{\mathbb{R}})$ and $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mathrm{Im}(\mu)); \overline{\mathbb{R}})$ then f is ***integrable*** with respect to μ.

(vi) If $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_j); \overline{\mathbb{R}})$, $j \in \{1, \dots, n\}$, f is ***integrable*** with respect to $\boldsymbol{\mu}$.

For a subset $I \subseteq \overline{\mathbb{R}}$ we denote the set of I-valued functions integrable with respect to μ (resp. $\boldsymbol{\mu}$) by $\mathsf{L}^{(1)}((X, \mathscr{A}, \mu); I)$ (resp. $\mathsf{L}^{(1)}((X, \mathscr{A}, \boldsymbol{\mu}); I)$), or simply by $\mathsf{L}^{(1)}(X; I)$ if $\mathscr{A}$ and μ (resp. $\boldsymbol{\mu}$) are understood. •

Since, by virtue of the Jordan decomposition, integration with respect to signed, complex, and vector measures boils down to integration with respect to positive measures as usual, one anticipates that many of the properties of the integral with respect to positive measures will carry over to signed, complex, and vector measures. Let us record some of these.

First we relate the integral of a function with the integral with respect to a measure to the integral with respect to the variation of the measure.

5.7.59 Proposition (Characterisation of integrals with respect to signed, complex, and vector measures) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and for* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*, the following statements hold:*

(i) if μ *is a signed or complex measure on* $\mathscr{A}$*, then* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, |\mu|); \overline{\mathbb{R}})$*, and if either of these equivalent statements holds, then*

$$\left|\int_{\mathsf{X}} \mathsf{f}\,\mathrm{d}\mu\right| \le \int_{\mathsf{X}} |\mathsf{f}|\,\mathrm{d}|\mu|;$$

(ii) if $\boldsymbol{\mu}$ *is a vector measure on* $\mathscr{A}$ *taking values in* $\mathbb{R}^{\mathsf{n}}$*, then* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *if and only if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \|\mu\|_{\mathbb{R}^{\mathsf{n}}}); \overline{\mathbb{R}})$*, and if either of these equivalent statements holds, then*

$$\left\|\int_{\mathsf{X}} \mathsf{f}\,\mathrm{d}\boldsymbol{\mu}\right\|_{\mathbb{R}^{\mathsf{n}}} \le \int_{\mathsf{X}} |\mathsf{f}|\,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^{\mathsf{n}}}.$$

Proof Let us first consider the case where μ is a signed measure on $\mathscr{A}$ with Jordan decomposition $\mu = \mu_+ - \mu_-$. If $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ then, by definition $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_+); \overline{\mathbb{R}})$ and $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_-); \overline{\mathbb{R}})$. Therefore,

$$\int_X |f|\,\mathrm{d}|\mu| = \int_X |f|\,\mathrm{d}\mu_+ + \int_X |f|\,\mathrm{d}\mu_- < \infty,$$

and so $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, |\mu|); \overline{\mathbb{R}})$. Conversely, suppose that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, |\mu|); \overline{\mathbb{R}})$. Then

$$\int_X |f|\,\mathrm{d}|\mu| = \int_X |f|\,\mathrm{d}\mu_+ + \int_X |f|\,\mathrm{d}\mu_- < \infty.$$

Thus $f \in \mathscr{L}^{(1)}((X, \mathscr{A}, \mu_+); \overline{\mathbb{R}})$ and $\in \mathscr{L}^{(1)}((X, \mathscr{A}, \mu_-); \overline{\mathbb{R}})$. Therefore,

$$\int_X f\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu_+ - \int_X f\,\mathrm{d}\mu_-$$

is well-defined, and so $f \in \mathscr{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$.

For the final assertion of this part of the theorem, we compute

$$\begin{aligned}\left|\int_X f\,\mathrm{d}\mu\right| &= \left|\int_X f\,\mathrm{d}\mu_+ - \int_X f\,\mathrm{d}\mu_-\right| \le \left|\int_X f\,\mathrm{d}\mu_+\right| + \left|\int_X f\,\mathrm{d}\mu_-\right| \\ &\le \int_X |f|\,\mathrm{d}\mu_+ + \int_X |f|\,\mathrm{d}\mu_- = \int_X |f|\,\mathrm{d}|\mu|,\end{aligned}$$

as claimed.

Now we consider the case of a vector measure $\boldsymbol{\mu}$, the case of a complex measure following from this as a special case. Suppose first that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ so that, by definition, $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_j); \overline{\mathbb{R}})$ for each $j \in \{1, \dots, n\}$. Let us first suppose that f is a nonnegative-valued simple function. Thus

$$f = \sum_{j=1}^{k} c_j \chi_{A_j}$$

for $c_j \in \overline{\mathbb{R}}_{\geq 0}$, $j \in \{1, \dots, k\}$, and for pairwise disjoint measurable sets A_j, $j \in \{1, \dots, k\}$. Then

$$\int_X f \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} = \sum_{j=1}^{k} c_j \|\boldsymbol{\mu}\|_{\mathbb{R}^n}(A_j) \leq \sum_{j=1}^{k} c_j \sum_{l=1}^{n} |\mu_l|(A_j),$$

the last inequality holding by (5.8). Noting that

$$\int_X f \,\mathrm{d}|\mu_l| = \sum_{j=1}^{k} c_j |\mu_l|(A_j),$$

we deduce that

$$\int_X f \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} \leq \sum_{l=1}^{n} \int_X f \mathrm{d}|\mu_l|,$$

giving $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$ in the case when f is a nonnegative simple function. For a general nonnegative function $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ we let $(f_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence of nonnegative simple functions such that $f_j(x) \leq f_{j+1}(x)$ for $x \in X$ and $j \in \mathbb{Z}_{>0}$ and such that $\lim_{j \to \infty} f_j(x) = f(x)$; see Proposition 5.6.39. Then

$$\int_X f_j \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} \leq \sum_{l=1}^{n} \int_X f_j \mathrm{d}|\mu_l| \leq \sum_{l=1}^{n} \int_X f \mathrm{d}|\mu_l|,$$

the last inequality by Proposition 5.7.19. Thus, by the Monotone Convergence Theorem,

$$\int_X f \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} = \lim_{j \to \infty} \int_X f_j \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} \leq \sum_{l=1}^{n} \int_X f \mathrm{d}|\mu_l|,$$

giving $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$ for a nonnegative $\boldsymbol{\mu}$-integrable function f. For a general $\boldsymbol{\mu}$-integrable function f we then have

$$\int_X |f| \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} = \lim_{j \to \infty} \int_X f_j \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} \leq \sum_{l=1}^{n} \int_X |f| \mathrm{d}|\mu_l|,$$

giving $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$.

Now we suppose that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$. As above, we first suppose that f is a nonnegative-valued simple function:

$$f = \sum_{j=1}^{k} c_j \chi_{A_j}.$$

For $l \in \{1, \dots, n\}$ we have

$$\begin{aligned}\int_X f \,\mathrm{d}|\mu_l| \leq \sum_{l=1}^n \int_X f \,\mathrm{d}|\mu_l| &\leq \sum_{l=1}^n \sum_{j=1}^k c_j |\mu_l|(A_j)\\ &\leq \sqrt{n} \sum_{j=1}^k c_j \|\boldsymbol{\mu}\|_{\mathbb{R}^n}(A_j) = \sqrt{n} \int_X f \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n},\end{aligned}$$

using Exercise 5.3.6 and Proposition 5.3.55. Thus, for nonnegative simple functions $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$ we have $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_l); \overline{\mathbb{R}})$, $l \in \{1, \dots, n\}$, and so $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$. Now one can prove that

$$\int_X f \,\mathrm{d}|\mu_l| \leq \sum_{l=1}^n \int_X f \,\mathrm{d}|\mu_l| \leq \sqrt{n} \int_X f \mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n}$$

for general nonnegative functions $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$ using an argument involving a sequence of simple functions a $(f_j)_{j \in \mathbb{Z}_{>0}}$ approximating f, just as in the preceding paragraph. Also just as in the preceding paragraph, it follows that, for a general $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^n}); \overline{\mathbb{R}})$,

$$\int_X |f| \,\mathrm{d}|\mu_l| \leq \sum_{l=1}^n \int_X |f| \,\mathrm{d}|\mu_l| \leq \sqrt{n} \int_X |f| \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n}, \tag{5.19}$$

and so $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$.

Moreover, by Proposition **??**, by the fact that the proposition holds for signed measures, and by (5.19), we have

$$\Big\| \int_X f \,\mathrm{d}\boldsymbol{\mu} \Big\|_{\mathbb{R}^n} \leq \sum_{l=1}^n \Big| \int_X f \,\mathrm{d}\mu_l \Big| \leq \sum_{l=1}^n \int_X |f| \,\mathrm{d}|\mu_l| \leq \sqrt{n} \int_X |f| \,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n},$$

which gives the final assertion of the proposition. ■

First we can show that the integral depends, in the appropriate sense, on the value of a function up to a set of measure zero.

5.7.60 Proposition (Integrals of functions agreeing almost everywhere) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and for* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$, *the following statements hold:*

(i) if μ *is a signed or complex measure on* $\mathscr{A}$ *and if*

$$|\mu|(\{\mathsf{x} \in \mathsf{X} \mid \mathsf{f}(\mathsf{x}) \neq \mathsf{g}(\mathsf{x})\}) = 0,$$

then $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and, if either of these conditions holds,*

$$\int_{\mathsf{X}} \mathsf{f} \,\mathrm{d}\mu = \int_{\mathsf{X}} \mathsf{g} \,\mathrm{d}\mu;$$

(ii) if $\boldsymbol{\mu}$ is a vector measure on $\mathscr{A}$ taking values in $\mathbb{R}^{\mathrm{n}}$ and if

$$\|\mu\|_{\mathbb{R}^{\mathrm{n}}}(\{\mathrm{x} \in \mathrm{X} \mid \ \mathrm{f}(\mathrm{x}) \neq \mathrm{g}(\mathrm{x})\}) = 0,$$

then $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *if and only if* $\mathrm{g} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *and, if either of these conditions holds,*

$$\int_{\mathrm{X}} \mathrm{f}\, \mathrm{d}\boldsymbol{\mu} = \int_{\mathrm{X}} \mathrm{g}\, \mathrm{d}\boldsymbol{\mu}.$$

Proof Let us first consider the case of a signed measure μ. First suppose that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$. By Proposition 5.7.59 it follows that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, |\mu|); \overline{\mathbb{R}})$. Since g differs from f on a set whose $|\mu|$-measure is zero, it follows from Proposition 5.7.11 that $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, |\mu|); \overline{\mathbb{R}})$ and so $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$, again by Proposition 5.7.59. Of course, the argument is reversible, showing that if $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ then $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$. If Z is the set of points where f and g differ, then

$$\Big|\int_Z (f - g)\, \mathrm{d}\mu\Big| \leq \int_Z |f - g|\, \mathrm{d}|\mu| = 0,$$

the first inequality by Proposition 5.7.59. Therefore, using the Proposition 5.7.62 below, we have

$$\int_X (f - g)\, \mathrm{d}\mu = \int_{X \setminus Z} (f - g)\, \mathrm{d}\mu + \int_Z (f - g)\, \mathrm{d}\mu = \int_{X \setminus Z} (f - g)\, \mathrm{d}\mu = 0.$$

By Proposition 5.7.61 we then have

$$\int_X f\, \mathrm{d}\mu = \int_X g\, \mathrm{d}\mu,$$

giving the first part of the result.

To conclude, we prove the proposition for vector measures, the complex case being a consequence of this. Suppose that Z denotes the set of points where f and g differ. Then

$$|\mu_l|(Z) = \int_X \chi_Z\, \mathrm{d}\mu_l \leq \sqrt{n} \int_X \chi_Z\, \mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n} = 0, \tag{5.20}$$

where we have used (5.19). Then the first part of the proof gives $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_l); \overline{\mathbb{R}})$ if and only if $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu_l); \overline{\mathbb{R}})$ for each $l \in \{1, \dots, n\}$. The definition of the integral with respect to $\boldsymbol{\mu}$, along with the conclusions from the first part of the result, gives

$$\int_X f\, \mathrm{d}\boldsymbol{\mu} = \int_X g\, \mathrm{d}\boldsymbol{\mu},$$

as desired. ■

The following result concerning algebraic operations can be deduced immediately by applying the corresponding result for positive measures to the Jordan decomposition of the measures involved.

5.7.61 Proposition (Algebraic operations for the integral with respect to signed, complex, and signed measures) *For a measurable space* $(\mathsf{X}, \mathscr{A})$*, the following statements hold:*

(i) *if* μ *is a signed or complex measure and if* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$*, then* $\mathsf{f} + \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and*

$$\int_{\mathsf{X}} (\mathsf{f} + \mathsf{g})\, \mathrm{d}\mu = \int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\mu + \int_{\mathsf{X}} \mathsf{g}\, \mathrm{d}\mu;$$

(ii) *if* $\boldsymbol{\mu}$ *is a vector measure taking values in* $\mathbb{R}^{\mathsf{n}}$ *and if* $\mathsf{f}, \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$*, then* $\mathsf{f} + \mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *and*

$$\int_{\mathsf{X}} (\mathsf{f} + \mathsf{g})\, \mathrm{d}\boldsymbol{\mu} = \int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\boldsymbol{\mu} + \int_{\mathsf{X}} \mathsf{g}\, \mathrm{d}\boldsymbol{\mu};$$

(iii) *if* μ *is a signed or complex measure, if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and if* $\alpha \in \mathbb{R}$*, then* $\alpha\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and*

$$\int_{\mathsf{X}} \alpha\mathsf{f}\, \mathrm{d}\mu = \alpha \int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\mu;$$

(iv) *if* $\boldsymbol{\mu}$ *is a vector measure taking values in* $\mathbb{R}^{\mathsf{n}}$*, if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$*, and if* $\alpha \in \mathbb{R}$*, then* $\alpha\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *and*

$$\int_{\mathsf{X}} \alpha\mathsf{f}\, \mathrm{d}\boldsymbol{\mu} = \alpha \int_{\mathsf{X}} \mathsf{f}\, \mathrm{d}\boldsymbol{\mu}.$$

Proof We first consider the case of a signed measure μ with Jordan decomposition $\mu = \mu_+ - \mu_-$. We then have

$$\begin{aligned}\int_X (f + g)\, \mathrm{d}\mu &= \int_X (f + g)\, \mathrm{d}\mu_+ - \int_X (f - g)\, \mathrm{d}\mu_- \\ &= \int_X f\, \mathrm{d}\mu_+ + \int_X g\, \mathrm{d}\mu_+ - \int_X f\, \mathrm{d}\mu_- - \int_X g\, \mathrm{d}\mu_- \\ &= \int_X f\, \mathrm{d}\mu + \int_X g\, \mathrm{d}\mu\end{aligned}$$

by Proposition 5.7.17. A similarly styled argument gives

$$\int_X \alpha f\, \mathrm{d}\mu = \alpha \int_X f\, \mathrm{d}\mu.$$

The result for vector measures then follows immediately from the result for signed measures by the definition of the integral with respect to a vector measure. The result for complex measures is a special case of the result for vector measures. ■

We can also break integrals with respect to signed, complex, and vector measures into separate integrals over disjoint sets.

5.7.62 Proposition (Breaking the integral in two) *For a measurable space* $(\mathrm{X}, \mathscr{A})$ *let* $\mathrm{A}, \mathrm{B} \in \mathscr{A}$ *be such that* $\mathrm{X} = \mathrm{A} \overset{\circ}{\cup} \mathrm{B}$ *and let* $\mathrm{f} \in \mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \overline{\mathbb{R}})$*. Then the following statements hold:*

(i) if μ *is a signed or complex measure on* $\mathscr{A}$*, then* $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $\mathrm{f}|\mathrm{A} \in \mathsf{L}^{(1)}((\mathrm{A}, \mathscr{A}_\mathrm{A}, \mu|\mathscr{A}_\mathrm{A}); \overline{\mathbb{R}})$ *and* $\mathrm{f}|\mathrm{B} \in \mathsf{L}^{(1)}((\mathrm{B}, \mathscr{A}_\mathrm{B}, \mu|\mathscr{A}_\mathrm{B}); \overline{\mathbb{R}})$*, and if either of these two equivalent conditions holds,*

$$\int_\mathrm{X} \mathrm{f}\,\mathrm{d}\mu = \int_\mathrm{A} (\mathrm{f}|\mathrm{A})\,\mathrm{d}\mu_\mathrm{A} + \int_\mathrm{B} (\mathrm{f}|\mathrm{B})\,\mathrm{d}\mu_\mathrm{B};$$

(ii) if $\boldsymbol{\mu}$ *is a vector measure on* $\mathscr{A}$*, then* $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *if and only if* $\mathrm{f}|\mathrm{A} \in \mathsf{L}^{(1)}((\mathrm{A}, \mathscr{A}_\mathrm{A}, \boldsymbol{\mu}|\mathscr{A}_\mathrm{A}); \overline{\mathbb{R}})$ *and* $\mathrm{f}|\mathrm{B} \in \mathsf{L}^{(1)}((\mathrm{B}, \mathscr{A}_\mathrm{B}, \boldsymbol{\mu}|\mathscr{A}_\mathrm{B}); \overline{\mathbb{R}})$*, and if either of these two equivalent conditions holds,*

$$\int_\mathrm{X} \mathrm{f}\,\mathrm{d}\boldsymbol{\mu} = \int_\mathrm{A} (\mathrm{f}|\mathrm{A})\,\mathrm{d}\boldsymbol{\mu}_\mathrm{A} + \int_\mathrm{B} (\mathrm{f}|\mathrm{B})\,\mathrm{d}\boldsymbol{\mu}_\mathrm{B};$$

Proof We first consider the case of a signed measure μ. By Proposition 5.7.22 it follows that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, |\mu|); \overline{\mathbb{R}})$ if and only if $f|A \in \mathsf{L}^{(1)}((A, \mathscr{A}_A, |\mu||\mathscr{A}_A); \overline{\mathbb{R}})$ and $f|B \in \mathsf{L}^{(1)}((B, \mathscr{A}_B, |\mu||\mathscr{A}_B); \overline{\mathbb{R}})$. By Proposition 5.7.59 it follows that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ if and only if $f|A \in \mathsf{L}^{(1)}((A, \mathscr{A}_A, \boldsymbol{\mu}|\mathscr{A}_A); \overline{\mathbb{R}})$ and $f|B \in \mathsf{L}^{(1)}((B, \mathscr{A}_B, \boldsymbol{\mu}|\mathscr{A}_B); \overline{\mathbb{R}})$, as claimed. Moreover, writing $f = f\chi_A + f\chi_B$, we use Proposition 5.7.61 to give

$$\int_X f\,\mathrm{d}\mu = \int_A (f|A)\,\mathrm{d}\mu_A + \int_B (f|B)\,\mathrm{d}\mu_B.$$

The result for vector and complex measures follows immediately from the conclusion for signed measures, using the definition of the integral in these cases. ■

5.7.63 Corollary (Breaking the integral almost in two) ***missing stuff*** *For a measurable space* $(\mathrm{X}, \mathscr{A})$ *let* $\mathrm{A}, \mathrm{B} \in \mathscr{A}$ *be such that* $\mathrm{X} = \mathrm{A} \cup \mathrm{B}$ *and let* $\mathrm{f} \in \mathsf{L}^{(0)}((\mathrm{X}, \mathscr{A}); \overline{\mathbb{R}})$*. Then the following statements hold:*

(i) if μ *is a signed or complex measure on* $\mathscr{A}$ *and if* $|\mu|(\mathrm{A}) = 0$*, then* $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *if and only if* $\mathrm{f}|\mathrm{A} \in \mathsf{L}^{(1)}((\mathrm{A}, \mathscr{A}_\mathrm{A}, \mu|\mathscr{A}_\mathrm{A}); \overline{\mathbb{R}})$ *and* $\mathrm{f}|\mathrm{B} \in \mathsf{L}^{(1)}((\mathrm{B}, \mathscr{A}_\mathrm{B}, \mu|\mathscr{A}_\mathrm{B}); \overline{\mathbb{R}})$*;*

(ii) if $\boldsymbol{\mu}$ *is a vector measure on* $\mathscr{A}$ *and if* $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}(\mathrm{A}) = 0$*, then* $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$ *if and only if* $\mathrm{f}|\mathrm{A} \in \mathsf{L}^{(1)}((\mathrm{A}, \mathscr{A}_\mathrm{A}, \boldsymbol{\mu}|\mathscr{A}_\mathrm{A}); \overline{\mathbb{R}})$ *and* $\mathrm{f}|\mathrm{B} \in \mathsf{L}^{(1)}((\mathrm{B}, \mathscr{A}_\mathrm{B}, \boldsymbol{\mu}|\mathscr{A}_\mathrm{B}); \overline{\mathbb{R}})$*.*

Proof This follows from Propositions 5.7.60 5.7.62. ■

Finally, for signed, complex, and vector measures we have a version of the Dominated Convergence Theorem. Note here that a little care must be exercised in stating the hypotheses.

5.7.64 Theorem (Dominated Convergence Theorem for signed, complex, and vector measures) *For a measurable space* $(\mathsf{X}, \mathscr{A})$ *and for a sequence* $(\mathsf{f}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ *in* $\mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *the following statements hold:*

(i) *if* μ *is a signed or complex measure and if*

(a) *the limit* $\mathrm{f}(\mathrm{x}) = \lim_{\mathrm{j}\to\infty} \mathrm{f}_\mathrm{j}(\mathrm{x})$ *exists for* $|\mu|$*-almost every* $\mathrm{x} \in \mathrm{X}$ *and if*

(b) *there exists* $\mathrm{g} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, |\mu|); \overline{\mathbb{R}}_{\geq 0})$ *such that, for* $|\mu|$*-almost every* $\mathrm{x} \in \mathrm{X}$, $|\mathrm{f}_\mathrm{j}|(\mathrm{x}) \leq \mathrm{g}(\mathrm{x})$ *for every* $\mathrm{j} \in \mathbb{Z}_{>0}$,

then $\mathrm{f}, \mathrm{f}_\mathrm{j} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$, $\mathrm{j} \in \mathbb{Z}_{>0}$, *and*

$$\int_\mathrm{X} \mathrm{f}\,\mathrm{d}\mu = \lim_{\mathrm{j}\to\infty} \int_\mathrm{X} \mathrm{f}_\mathrm{j}\,\mathrm{d}\mu;$$

(ii) *if* $\boldsymbol{\mu}$ *is a vector measure taking values in* $\mathbb{R}^\mathrm{n}$ *and if*

(a) *the limit* $\mathrm{f}(\mathrm{x}) = \lim_{\mathrm{j}\to\infty} \mathrm{f}_\mathrm{j}(\mathrm{x})$ *exists for* $\|\boldsymbol{\mu}\|_{\mathbb{R}^\mathrm{n}}$*-almost every* $\mathrm{x} \in \mathrm{X}$ *and if*

(b) *there exists* $\mathrm{g} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \|\boldsymbol{\mu}\|_{\mathbb{R}^\mathrm{n}}); \overline{\mathbb{R}}_{\geq 0})$ *such that, for* $\|\boldsymbol{\mu}\|_{\mathbb{R}^\mathrm{n}}$*-almost every* $\mathrm{x} \in \mathrm{X}$, $|\mathrm{f}_\mathrm{j}|(\mathrm{x}) \leq \mathrm{g}(\mathrm{x})$ *for every* $\mathrm{j} \in \mathbb{Z}_{>0}$,

then $\mathrm{f}, \mathrm{f}_\mathrm{j} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \boldsymbol{\mu}); \overline{\mathbb{R}})$, $\mathrm{j} \in \mathbb{Z}_{>0}$, *and*

$$\int_\mathrm{X} \mathrm{f}\,\mathrm{d}\boldsymbol{\mu} = \lim_{\mathrm{j}\to\infty} \int_\mathrm{X} \mathrm{f}_\mathrm{j}\,\mathrm{d}\boldsymbol{\mu}.$$

Proof We first consider the case of a signed measure μ with Jordan decomposition $\mu = \mu_+ - \mu_-$. The integrability of f and f_j, $j \in \mathbb{Z}_{>0}$, with respect to μ follows from their assumed integrability with respect to $|\mu|$, along with Proposition 5.7.59. Since $|\mu| = \mu_+ + \mu_-$, it follows that the limit $f(x) = \lim_{j\to\infty} f_j(x)$ exists for μ_+-almost every $x \in X$ and for μ_--almost every $x \in X$. Also, $|\mu|$-integrability of g implies μ_+- and μ_--integrability of g. Finally, we have $|f_j|(x) \leq g(x)$ for μ_+- and μ_--almost every $x \in X$ and for every $j \in \mathbb{Z}_{>0}$. Then we compute

$$\begin{aligned}\lim_{j\to\infty} \int_X f_j\,\mathrm{d}\mu &= \lim_{j\to\infty}\Big(\int_X f_j\,\mu_+ - \int_X f_j\,\mathrm{d}\mu_-\Big)\\ &= \lim_{j\to\infty} \int_X f_j\,\mu_+ - \lim_{j\to\infty} \int_X f_j\,\mathrm{d}\mu_-\\ &= \int_X f\,\mathrm{d}\mu_+ - \int_X f\,\mathrm{d}\mu_- = \int_X f\,\mathrm{d}\mu,\end{aligned}$$

using the Dominated Convergence Theorem for positive measures, along with the commutativity of limits with sums (Proposition 2.3.23).

We next prove the theorem for the case of a vector measure, noting that the case of complex measures follows from this. As in (5.20), if Z has $\|\boldsymbol{\mu}\|_{\mathbb{R}^n}$-measure zero, then Z also has $|\mu_l|$-measure zero for each $l \in \{1, \dots, n\}$. Therefore, the hypotheses of the theorem give:

1. the limit $f(x) = \lim_{j\to\infty} f_j(x)$ exists for $|\mu_l|$-almost $x \in X$ for each $l \in \{1, \dots, l\}$;
2. $|f_j|(x) \leq g(x)$ for $|\mu_l|$-almost every $x \in X$ for each $j \in \mathbb{Z}_{>0}$ and $l \in \{1, \dots, n\}$.

As we saw in the proof of Proposition 5.7.59,

$$\int_X g\,\mathrm{d}|\mu_l| \leq \sqrt{n}\int_X g\,\mathrm{d}\|\boldsymbol{\mu}\|_{\mathbb{R}^n},$$

and so our hypotheses imply that $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, |\mu_l|); \overline{\mathbb{R}}_{\geq 0})$ for each $l \in \mathbb{Z}_{>0}$. This all implies that the result from the first part of the theorem gives the result for vector measures. ■

We next show how signed, complex, and vector measures can be built from positive measures and integrable functions. This gives us a wealth of signed, complex, and vector measures. We shall see in ***missing stuff***, moreover, that an important class of measures arise *exactly* in the manner of the next result.

5.7.65 Proposition (Signed, complex, and vector measures from functions) *If* $(\mathrm{X}, \mathscr{A}, \mu)$ *is a measure space, then the following statements hold:*

(i) if $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \mathbb{R})$ *then* $\mathrm{f}\cdot\mu\colon \mathscr{A} \to \overline{\mathbb{R}}$ *defined by*

$$(\mathrm{f}\cdot\mu)(\mathrm{A}) = \int_{\mathrm{X}} \mathrm{f}\chi_{\mathrm{A}}\,\mathrm{d}\mu$$

is a finite signed measure on $\mathscr{A}$*;*

(ii) if $\mathrm{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \mathbb{C})$ *then* $\mathrm{f}\cdot\mu\colon \mathscr{A} \to \mathbb{C}$ *defined by*

$$(\mathrm{f}\cdot\mu)(\mathrm{A}) = \int_{\mathrm{X}} \mathrm{f}\chi_{\mathrm{A}}\,\mathrm{d}\mu$$

is a complex measure on $\mathscr{A}$*;*

(iii) if $\mathbf{f} \in \mathsf{L}^{(1)}((\mathrm{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathrm{n}})$ *then* $\mathbf{f}\cdot\mu\colon \mathscr{A} \to \mathbb{R}^{\mathrm{n}}$ *defined by*

$$(\mathbf{f}\cdot\mu)(\mathrm{A}) = \int_{\mathrm{X}} \mathbf{f}\chi_{\mathrm{A}}\,\mathrm{d}\mu$$

is a vector measure on $\mathscr{A}$*.*

Proof We prove the statement for vector measures, since the other cases are a special case of this.

It is clear that $(\boldsymbol{f}\cdot\mu)(\emptyset) = 0$. Now let $(A_j)_{j\in\mathbb{Z}_{>0}}$ be a family of pairwise disjoint elements of $\mathscr{A}$ and let $A = \cup_{j\in\mathbb{Z}_{>0}} A_j$. If $g = \|\boldsymbol{f}\|_{\mathbb{R}^n}\chi_A$ then $g(x) \leq \|\boldsymbol{f}\|_{\mathbb{R}^n}(x)$ for every $x \in X$ and so $g \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}_{\geq 0})$ by Proposition 5.7.47. If we define $B_k = \cup_{j=1}^k A_j$ and $\boldsymbol{f}_k = \boldsymbol{f}\chi_k$, $k \in \mathbb{Z}_{>0}$, then

$$\lim_{k\to\infty} \boldsymbol{f}_k(x) = \boldsymbol{f}(x)\chi_A(x), \qquad x \in X.$$

Therefore, by the Dominated Convergence Theorem, Theorem 5.7.56,

$$(\boldsymbol{f}\cdot\mu)(A) = \int_X \boldsymbol{f}\chi_A\,\mathrm{d}\mu = \lim_{k\to\infty}\int_X \boldsymbol{f}_k\,\mathrm{d}\mu = \lim_{k\to\infty}\sum_{j=1}^k\int_X \boldsymbol{f}\chi_{A_j}\,\mathrm{d}\mu = \sum_{j=1}^\infty(\boldsymbol{f}\cdot\mu)(A_j),$$

giving countable additivity of $\boldsymbol{f}\cdot\mu$. ■

For the measures determined by integrable functions, as in Proposition 5.7.65, it is possible to explicitly characterise the integrals with respect to these measures. The notation from the previous proposition will be used in the statement of the next.

5.7.66 Proposition (Integration with respect to measures from functions) *If* $(\mathsf{X}, \mathscr{A}, \mu)$ *is a measure space and if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$, $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{C})$, *and* $\mathbf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathsf{n}})$, *then the following statements hold:*

(i) if $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R})$ *then* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mathsf{f} \cdot \mu); \mathbb{R})$ *if and only if* $\mathsf{fg} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R})$, *and if either of these equivalent conditions holds,*

$$\int_{\mathsf{X}} \mathsf{g}\, \mathrm{d}(\mathsf{f} \cdot \mu) = \int_{\mathsf{X}} (\mathsf{fg})\, \mathrm{d}\mu;$$

(ii) if $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{C})$ *then* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mathsf{f} \cdot \mu); \mathbb{R})$ *if and only if* $\mathsf{fg} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{C})$, *and if either of these equivalent conditions holds,*

$$\int_{\mathsf{X}} \mathsf{g}\, \mathrm{d}(\mathsf{f} \cdot \mu) = \int_{\mathsf{X}} (\mathsf{fg})\, \mathrm{d}\mu;$$

(iii) if $\mathbf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathsf{n}})$ *then* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mathbf{f} \cdot \mu); \mathbb{R})$ *if and only if* $\mathsf{g}\mathbf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathsf{n}})$, *and if either of these equivalent conditions holds,*

$$\int_{\mathsf{X}} \mathsf{g}\, \mathrm{d}(\mathbf{f} \cdot \mu) = \int_{\mathsf{X}} (\mathsf{g}\mathbf{f})\, \mathrm{d}\mu.$$

Proof Let us first consider the case where $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R})$. Let us define

$$P = \{x \in X \mid f(x) \geq 0\}, \quad N = X \setminus P,$$

noting that P (and so N) is measurable by Proposition 5.6.16. Clearly (P, N) is a Hahn decomposition for $(X, \mathscr{A}, f \cdot \mu)$. Moreover, the corresponding Jordan decomposition is

$$f \cdot \mu = f_+ \cdot \mu - f_- \cdot \mu,$$

where, as usual, $f_+(x) = \max\{f(x), 0\}$ and $f_-(x) = \max\{-f(x), 0\}$. Noting that gf is integrable if and only if both gf_+ and gf_- are integrable, and computing

$$\int_X (fg)\, \mathrm{d}\mu = \int_X (f_+ g)\, \mathrm{d}\mu - \int_X (f_- g)\, \mathrm{d}\mu = \int_X g\, \mathrm{d}(f_+ \cdot \mu) - \int_X g\, \mathrm{d}(f_- \cdot \mu) = \int_X g\, \mathrm{d}(f \cdot \mu),$$

the result for signed measures follows.

To complete the proof, we suppose that $\boldsymbol{f} \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \mathbb{R}^n)$, and prove the last assertion in the statement of the proposition. The proof of the second assertion is a consequence of this. For $A \in \mathscr{A}$ and $l \in \{1, \dots, n\}$ we have

$$(\boldsymbol{f} \cdot \mu)_l(A) = \mathrm{pr}_l\Big(\int_X \boldsymbol{f} \chi_A\, \mathrm{d}\mu\Big) = \int_X f_l \chi_A\, \mathrm{d}\mu = (f_l \cdot \mu)(A),$$

where $\mathrm{pr}_l \colon \mathbb{R}^n \to \mathbb{R}$ is the projection onto the lth component. Given this, and the definitions of the integral with respect to a vector measure and the integral of a vector-valued function, the result follows from the result proved above for $\mathbb{R}$-valued functions. ■

5.7.9 Notes

There is no standard convention on what Beppo Levi's Theorem is. Sometimes what we call the Monotone Convergence Theorem is called Beppo Levi's Theorem.

Exercises

5.7.1 Let $(X, \mathscr{A}, \mu)$ be a measure space and let $f, g \in \mathsf{S}(X; \overline{\mathbb{R}}_{\geq 0})$ satisfy $f(x) \leq g(x)$ for each $x \in X$. Show that

$$\int_X f \, \mathrm{d}\mu \leq \int_X g \, \mathrm{d}\mu.$$

5.7.2 Let $(X, \mathscr{A}, \mu)$ be a measure space and let $f \in \mathsf{S}(X; \overline{\mathbb{R}})$. For $A \in \mathscr{A}$ define $f_A \colon X \to \overline{\mathbb{R}}$ by $f_A = f\chi_A$. Show that

$$\int_X f_A \, \mathrm{d}\mu = \int_A (f|A) \, \mathrm{d}\mu_A.$$

5.7.3 Let $X = \mathbb{Z}_{>0}$, let $\mathscr{A} = \mathbf{2}^{\mathbb{Z}_{>0}}$, and let $\mu_\Sigma \colon \mathscr{A} \to \overline{\mathbb{R}}_{\geq 0}$ be the counting measure:

$$\mu_\Sigma(A) = \begin{cases} \mathrm{card}(A), & \mathrm{card}(A) < \infty, \\ \infty, & \text{otherwise.} \end{cases}$$

Verify the following statements using only the definition of the integral, i.e., do not use the general constructions of Examples 5.7.7 and 5.7.10.

(a) A function $f \colon \mathbb{Z}_{>0} \to \mathbb{R}$ is integrable if and only if the series $\sum_{j=1}^{\infty} f(j)$ is absolutely convergent.

(b) If f is integrable then

$$\int_{\mathbb{Z}_{>0}} f \, \mathrm{d}\mu_\Sigma = \sum_{j=1}^{\infty} f(j).$$

5.7.4 For a measure space $(X, \mathscr{A}, \mu)$ and for $f \in \mathsf{L}^{(0)}((X, \mathscr{A}); \overline{\mathbb{R}})$, show that $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$ if and only if $|f| \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu); \overline{\mathbb{R}})$.

5.7.5 Let $X = \mathbb{Z}_{>0}$, $\mathscr{A} = \mathbf{2}^X$, and let μ_Σ be the counting measure on $\mathscr{A}$. Define $f \colon \mathbb{Z}_{>0} \to \mathbb{R}$ by $f(j) = j$. Use the Monotone Convergence Theorem to show that $f \notin \mathsf{L}^{(1)}((\mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}}, \mu_\Sigma); \mathbb{R})$.

The following exercise requires the notion of the concept of a norm which will be introduced in Section 6.1.

5.7.6 Let $(X, \mathscr{A}, \mu)$ be a measure space, let $f \in \mathsf{L}^{(1)}((X, \mathscr{A}, \mu), \mathbb{R}^n)$, and let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. Show that

$$\left\| \int_X f \, \mathrm{d}\mu \right\| \leq \int_X \|f\| \, \mathrm{d}\mu.$$

Hint: *Use Proposition 5.7.47 and Theorem 6.1.15.* ***missing stuff***

Section 5.8

Integration on products

In Section **??** we presented Fubini's Theorem for the Riemann integral which showed how the n-dimensional Riemann integral could be computed by means of one-dimensional integrals. In Section 5.3.6 we introduced the product measure on a finite product of measure spaces. Understanding these two things, it is then naturally ask whether the integral for a product measure can be understood in terms of the measure of the component measure spaces. The result is the general version of Fubini's Theorem. As part of our treatment of Fubini's Theorem, we give an alternative characterisation of the product measure.

Do I need to read this section? We shall make frequent use of Fubini's Theorem. That being said, to make use of Fubini's Theorem it is not necessary to understand all of the details we present here. What is most important is to understand the hypotheses of Fubini's Theorem. •

5.8.1 The product measure by integration

In Section 5.3.6 we defined a unique measure on a product of measure spaces that had a natural property in terms of the measure of measurable rectangles. In this section we retrieve this measure in another way, using the integral. This construction has the benefit of being simpler than that in Section 5.3.6, but only after one has the integral at hand.

In Section 5.3.6 we defined product measures for arbitrary finite products. However, it is notationally easier to deal with a product with two factors, and then use induction to arrive at the general case. Thus we consider two measure spaces $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$. As in Section 5.2.3, a ***measurable rectangle*** is a subset $A \times B \subseteq X \times Y$ where $A \in \mathscr{A}$ and $B \in \mathscr{B}$. We denote by $\sigma(\mathscr{A} \times \mathscr{B})$ the σ-algebra generated by the collection of measurable rectangles. For a set $E \subseteq X \times Y$ and for $(x, y) \in X \times Y$ we define subsets $E_x \subseteq Y$ and $E^y \subseteq X$ by

$$E_x = \{y' \in Y \mid (x, y') \in E\}, \qquad E^y = \{x' \in X \mid (x', y) \in E\}.$$

One calls the sets E_x and E^y ***sections*** of the set E.

The following result begins our construction of the product measure using the integral. The reader will hopefully recognise something Fubini-like in this result.

5.8.1 Lemma (Integrals of sections) *For σ-finite measure spaces $(\mathrm{X}, \mathscr{A}, \mu)$ and $(\mathrm{Y}, \mathscr{B}, \nu)$ and for $\mathrm{E} \in \sigma(\mathscr{A} \times \mathscr{B})$, define*

$$\begin{aligned} \phi_{\mathrm{E}} \colon \mathrm{X} &\to \overline{\mathbb{R}} \\ \mathrm{x} &\mapsto \nu(\mathrm{E_x}), \end{aligned} \qquad \begin{aligned} \psi_{\mathrm{E}} \colon \mathrm{Y} &\to \overline{\mathbb{R}} \\ \mathrm{y} &\mapsto \mu(\mathrm{E^y}). \end{aligned}$$

Then ϕ_{E} and ψ_{E} are $\mathscr{A}$-measurable and $\mathscr{B}$-measurable, respectively. Moreover,

$$\int_{\mathrm{X}} \phi_{\mathrm{E}}\,\mathrm{d}\mu = \int_{\mathrm{Y}} \psi_{\mathrm{E}}\,\mathrm{d}\nu.$$

Proof Denote by $\mathscr{M}(X \times Y)$ the collection of all sets E for which the conclusions of the lemma hold. We shall show that $\mathscr{M}(X \times Y)$ is a monotone class containing the set of measurable rectangles.

For $A \in \mathscr{A}$ and $B \in \mathscr{B}$ we have

$$\phi_{A\times B}(x) = \nu(B)\chi_A(x), \quad \psi_{A\times B}(y) = \mu(A)\chi_B(y),$$

which shows that $A \times B \in \mathscr{M}(X \times Y)$. Therefore, $\phi_{A\times B}$ and $\psi_{A\times B}$ are measurable (by Example 5.6.8–2) and

$$\int_X \phi_{A\times B}\,\mathrm{d}\mu = \int_Y \psi_{A\times B}\,\mathrm{d}\nu = \mu(A)\nu(B).$$

Thus $\mathscr{M}(X \times Y)$ contains the measurable rectangles.

Now let $(E_j)_{j\in\mathbb{Z}_{>0}}$ be a collection of subsets of $\mathscr{M}(A\times B)$ for which $E_j \subseteq E_{j+1}$, $j \in \mathbb{Z}_{>0}$. Then, denoting $E = \cup_{j\in\mathbb{Z}_{>0}} E_j$,

$$\lim_{j\to\infty} \phi_{E_j}(x) = \phi_E(x), \quad \lim_{j\to\infty} \psi_{E_j}(y) = \psi_E(y).$$

Thus $\phi_E \in \mathsf{L}^{(0)}((X,\mathscr{A}),\overline{\mathbb{R}})$ and $\psi_E \in \mathsf{L}^{(0)}((Y,\mathscr{B});\overline{\mathbb{R}})$ by Proposition 5.6.18. Note that the sequences $(\phi_{E_j}(x))_{j\in\mathbb{Z}_{>0}}$ and $(\psi_{E_j}(x))_{j\in\mathbb{Z}_{>0}}$ are monotonically increasing, so the Monotone Convergence Theorem gives

$$\int_X \phi_E\,\mathrm{d}\mu = \lim_{j\to\infty}\int_X \phi_{E_j}\,\mathrm{d}\mu = \lim_{j\to\infty}\int_Y \psi_{E_j}\,\mathrm{d}\nu = \int_Y \psi_E\,\mathrm{d}\nu.$$

Therefore, $E \in \mathscr{M}(X \times Y)$, which is part (i) of the definition of a monotone class.

Now, for the moment, suppose that $\mu(X)$ and $\nu(Y)$ are finite. Let $(E_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\sigma(\mathscr{A} \times \mathscr{B})$ such that $E_j \supseteq E_{j+1}$, $j \in \mathbb{Z}_{>0}$. Define $E = \cap_{j\in\mathbb{Z}_{>0}} E_j$ and note that

$$\lim_{j\to\infty} \phi_{E_j}(x) = \phi_E(x), \quad \lim_{j\to\infty} \psi_{E_j}(y) = \psi_E(y).$$

Thus $\phi_E \in \mathsf{L}^{(0)}((X,\mathscr{A}),\overline{\mathbb{R}})$ and $\psi_E \in \mathsf{L}^{(0)}((Y,\mathscr{B});\overline{\mathbb{R}})$ by Proposition 5.6.18. Note that we obviously have

$$\phi_{E_j}(x) \le \nu(Y)\chi_X(x),\ \phi_E(x) \le \nu(Y)\chi_X(x), \qquad \psi_{E_j}(y) \le \mu(X)\chi_Y(y),\ \psi_E(y) \le \mu(X)\chi_Y(y)$$

for every $(x, y) \in X, Y$. Moreover, since we are assuming that X and Y have finite measure we have $\chi_X \in \mathsf{L}^{(1)}((X,\mathscr{A},\mu);\mathbb{R})$ and $\chi_Y \in \mathsf{L}^{(1)}((Y,\mathscr{B},\nu);\mathbb{R})$. Therefore, the hypotheses of the Dominated Convergence Theorem hold and we have

$$\int_X \phi_E\,\mathrm{d}\mu = \lim_{j\to\infty}\int_X \phi_{E_j}\,\mathrm{d}\mu = \lim_{j\to\infty}\int_Y \psi_{E_j}\,\mathrm{d}\nu = \int_Y \psi_E\,\mathrm{d}\nu,$$

from which we conclude that $E \in \mathscr{M}(X \times X)$. This verifies part (ii) of Definition 5.2.11 in this case. Thus this shows that, when $\mu(X), \nu(Y) < \infty$, $\mathscr{M}(X \times Y)$ is a monotone class containing the measurable rectangles. From Theorem 5.2.13 it then follows that $\sigma(\mathscr{A}\times\mathscr{B})\subseteq\mathscr{M}(X\times Y)$. Thus the lemma holds in this case.

Now let us suppose that $\mu(X)$ and $\nu(Y)$ are not necessarily finite, but that using our assumption of σ-additivity we can write $X=\cup_{k\in\mathbb{Z}_{>0}}X_k$ and $Y=\cup_{k\in\mathbb{Z}_{>0}}Y_k$ where $\mu(X_k),\nu(Y_k)<\infty$, $k\in\mathbb{Z}_{>0}$, and where $(X_k)_{k\in\mathbb{Z}_{>0}}$ and $(Y_k)_{k\in\mathbb{Z}_{>0}}$ are pairwise disjoint measurable sets. Thus $X\times Y$ is the disjoint union of the measurable rectangles $X_{\times}Y_l$, $k,l\in\mathbb{Z}_{>0}$. Let $f\colon\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}\times\mathbb{Z}_{>0}$ be a bijection and, for $m\in\mathbb{Z}_{>0}$, define $Z_m=X_k\times Y_l$ where $\phi(m)=(k,l)$. Now $X\times Y$ is a disjoint union of the measurable sets Z_m, $m\in\mathbb{Z}_{>0}$. Finally, define $S_n=\cup_{m=1}^n Z_m$ so that $X\times Y$ is a union of the measurable sets S_n, $n\in\mathbb{Z}_{>0}$, where $S_n\subseteq S_{n+1}$. Note that $\mu(S_n)<\infty$ for every $n\in\mathbb{Z}_{>0}$ since S_n is a finite union of sets of finite measure.

Now let $E\in\sigma(\mathscr{A}\times\mathscr{B})$ and denote $E_n=E\cap S_n$. From our argument above, $E_n\in\sigma(\mathscr{A}\times\mathscr{B})$ and

$$\int_X\phi_{E_n}\,\mathrm{d}\mu=\int_Y\psi_{E_n}\,\mathrm{d}\nu.$$

We also have

$$\lim_{n\to\infty}\phi_{E_n}(x)=\phi_E(x),\quad \lim_{n\to\infty}\psi_{E_n}(y)=\psi_E(y)$$

for every $(x,y)\in X\times Y$. Since $S_n\subseteq S_{n+1}$ for every $n\in\mathbb{Z}_{>0}$, the sequences $(\phi_{E_n}(x))_{n\in\mathbb{Z}_{>0}}$ and $(\psi_{E_n}(y))_{n\in\mathbb{Z}_{>0}}$ are increasing for every $(x,y)\in X\times Y$. Therefore, by the Monotone Convergence Theorem,

$$\int_X\phi_E\,\mathrm{d}\mu=\lim_{n\to\infty}\int_X\phi_{E_n}\,\mathrm{d}\mu=\lim_{n\to\infty}\int_Y\psi_{E_n}\,\mathrm{d}\nu=\int_Y\psi_E\,\mathrm{d}\nu,$$

giving the lemma. ■

With the preceding, we can fairly easily derive the product measure using the integral.

5.8.2 Theorem (The product measure using the integral) *For σ-finite measure spaces $(\mathrm{X},\mathscr{A},\mu)$ and $(\mathrm{Y},\mathscr{B},\nu)$, the map $\mu\times\nu\colon\sigma(\mathscr{A}\times\mathscr{B})\to\overline{\mathbb{R}}_{\geq 0}$ defined by*

$$\mu\times\nu(\mathrm{E})=\int_{\mathrm{X}}\phi_{\mathrm{E}}\,\mathrm{d}\mu=\int_{\mathrm{Y}}\psi_{\mathrm{E}}\,\mathrm{d}\nu$$

makes $(\mathrm{X}\times\mathrm{Y},\sigma(\mathscr{A}\times\mathscr{B}),\mu\times\nu)$ a σ-finite measure space. Moreover, the measure $\mu\times\nu$ is the product measure as defined in Definition 5.3.34.

Proof It is clear that $\mu\times\lambda(\emptyset)=0$ since $\emptyset=\emptyset\times\emptyset$ is a measurable rectangle, being the product of two sets with zero measure. For a sequence $(E_j)_{j\in\mathbb{Z}_{>0}}$ of disjoint subsets of $\sigma(\mathscr{A}\times\mathscr{B})$ define $E=\cup_{j\in\mathbb{Z}_{>0}}E_j$. Note that

$$\phi_E(x)=\sum_{j=1}^\infty\phi_{E_j}(x),$$

and so Beppo Levi's Theorem gives

$$\mu\times\nu(E)=\int_X\phi_E\,\mathrm{d}\mu=\sum_{j=1}^\infty\int_X\phi_{E_j}\,\mathrm{d}\mu=\sum_{j=1}^\infty\mu\times\nu(E_j),$$

as desired.

That $\mu \times \nu$ is the product measure follows from Theorem 5.3.33, along with the fact that we showed in the proof of Lemma 5.8.1 that $\mu \times \nu(A \times B) = \mu(A)\nu(B)$ for $A \in \mathscr{A}$ and $B \in \mathscr{B}$. ■

Now that we have established the product measure using the integral for a product with two factors, it is more or less a straightforward induction to do the same for products with three or more factors. Indeed, suppose we have σ-finite measure spaces $(X_j, \mathscr{A}_j, \mu_j)$, $j \in \{1, \dots, k\}$. For $E \subseteq X_1 \times \dots \times X_k$ and for $x_k \in X_k$, denote

$$E_{x_k} = \{(x_1, \dots, x_{k-1}) \in X_1 \times \dots \times X_{k-1} \mid (x_1, \dots, x_{k-1}, x_k) \in E\}.$$

Suppose that we have defined the product measure $\mu_1 \times \dots \times \mu_{k-1}$ on $X_1 \times \dots \times X_{k-1}$. Then define $\phi_E \colon X_k \to \overline{\mathbb{R}}$ by

$$\phi_E(x_k) = \mu_1 \times \dots \times \mu_{k-1}(E_{x_k}).$$

We then have

$$\mu_1 \times \dots \times \mu_k(E) = \int_{X_k} \phi_{E_k} \, \mathrm{d}\mu_k,$$

which is the product measure.

5.8.2 The integral on product spaces

Either by the construction of the previous section, or by the construction of Section 5.3.6, we have defined on the product $X_1 \times \dots \times X_k$, for measure spaces $(X_j, \mathscr{A}_j, \mu_j)$, $j \in \{1, \dots, k\}$, a natural measure. One can then apply the construction of the integral from Section 5.7 to define the integral of measurable functions on the product. There is a slight hitch here that one needs to account for if one is to use this theory for the n-dimensional Lebesgue integral. To wit, in Section 5.5.4 we observed that the n-dimensional Lebesgue measure is not the product of the 1-dimensional Lebesgue measures on $\mathbb{R} \times \dots \times \mathbb{R}$, but is the completion of this measure. Thus we should develop integration for, not just the product measure, but its completion. This is not particularly difficult, but just requires a few additional words.

As in the preceding section, for simplicity we start with two measure spaces $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$. As in the preceding section, we denote by $\sigma(\mathscr{A} \times \mathscr{B})$ the natural product σ-algebra on $X \times Y$, i.e., the σ-algebra generated by the measurable rectangles. By $\mu \times \nu$ we denote the product measure. As we saw in Section 5.5.4 (and more generally in Remark 5.3.36), there are cases where the measure $\mu \times \nu$ is not complete (although there are also cases where the product measure *is* complete). Thus we denote by $(X \times Y, \overline{\sigma}(\mathscr{A} \times \mathscr{B}), \overline{\mu \times \nu})$ the completion of $(X \times Y, \sigma(\mathscr{A} \times \mathscr{B}), \mu \times \nu)$.

In the previous section we defined the notion of the sections for a subset $E \subseteq X \times Y$. This can also be done for functions. For a function $f \colon X \times Y \to \overline{\mathbb{R}}$, we define functions $f_x \colon Y \to \overline{\mathbb{R}}$ and $f^y \colon X \to \overline{\mathbb{R}}$ by

$$f_x(y) = f^y(x) = f(x, y).$$

One calls the functions f_x and f^y ***sections*** of the function f. The following result give the measurability properties of sections of sets and functions.

5.8.3 Lemma (Measurability of sections) *For measure spaces* $(\mathsf{X}, \mathscr{A}, \mu)$ *and* $(\mathsf{Y}, \mathscr{B}, \nu)$*, the following statements hold:*

(i) if $\mathsf{E} \subseteq \mathsf{X} \times \mathsf{Y}$ *is* $\sigma(\mathscr{A} \times \mathscr{B})$*-measurable, then* $\mathsf{E}_\mathsf{x} \in \mathscr{B}$ *for every* $\mathsf{x} \in \mathsf{X}$ *and* $\mathsf{E}^\mathsf{y} \in \mathscr{A}$ *for every* $\mathsf{y} \in \mathsf{Y}$*;*

(ii) if $(\mathsf{X}, \mathscr{A}, \mu)$ *and* $(\mathsf{Y}, \mathscr{B}, \nu)$ *are complete and if* $\mathsf{E} \subseteq \mathsf{X} \times \mathsf{Y}$ *is* $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$*-measurable, then* $\mathsf{E}_\mathsf{x} \in \mathscr{B}$ *for every* $\mathsf{x} \in \mathsf{X}$ *and* $\mathsf{E}^\mathsf{y} \in \mathscr{A}$ *for every* $\mathsf{y} \in \mathsf{Y}$*;*

(iii) if $\mathsf{f}\colon \mathsf{X} \times \mathsf{Y} \to \overline{\mathbb{R}}$ *is* $\sigma(\mathscr{A} \times \mathscr{B})$*-measurable, then* $\mathsf{f}_\mathsf{x} \in \mathsf{L}^{(0)}((\mathsf{Y}, \mathscr{B}); \overline{\mathbb{R}})$ *for almost every* $\mathsf{x} \in \mathsf{X}$ *and* $\mathsf{f}^\mathsf{y} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *for almost every* $\mathsf{y} \in \mathsf{Y}$*;*

(iv) if $(\mathsf{X}, \mathscr{A}, \mu)$ *and* $(\mathsf{Y}, \mathscr{B}, \nu)$ *are complete and if* $\mathsf{f}\colon \mathsf{X}\times\mathsf{Y} \to \overline{\mathbb{R}}$ *is* $\overline{\sigma}(\mathscr{A}\times\mathscr{B})$*-measurable, then* $\mathsf{f}_\mathsf{x} \in \mathsf{L}^{(0)}((\mathsf{Y}\,\mathrm{sB}); \overline{\mathbb{R}})$ *and* $\mathsf{f}^\mathsf{y} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$*.*

Proof (i) This is a special case of Proposition 5.2.18.

(ii) Next suppose that $E \in \overline{\sigma}(\mathscr{A} \times \mathscr{B})$. We let $U \subseteq E \subseteq L$ have the property that $U, L \in \sigma(\mathscr{A} \times \mathscr{B})$ and $\mu \times \nu(U \setminus L) = 0$. We may apply the first part of the proof to U to assert that U_x and L_x are measurable for all $x \in X$. Since $(U_x \setminus E_x) \subseteq (U_x \setminus L_x)$ and since $U_x \setminus L_x$ has measure zero, it follows that $U_x \setminus E_x$ is measurable by completeness of $\mathscr{A}$. Thus E_x is $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$-measurable. Similarly, E^y is also $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$-measurable.

(iii) Note that for $S \subseteq \mathbb{R}$ we have $f_x^{-1}(S) = (f^{-1}(S))_x$ and $(f^y)^{-1}(S) = (f^{-1}(S))^y$. This part of the lemma now follows from part (i).

(iv) By Proposition 5.7.15 we may find g that is $\sigma(\mathscr{A} \times \mathscr{B})$-measurable and for which $f(x, y) = g(x, y)$ except on a set that has zero measure relative to $\mu \times \nu$. Thus $h = f - g$ is zero except on a set that has zero measure relative to $\mu \times \nu$. This part of the lemma will follow from part (iii) if we can show that h_x and h^y are measurable for almost every $x \in X$ and $y \in Y$. If E is the set of points in $X \times Y$ where h does not vanish then $E \in \overline{\sigma}(\mathscr{A} \times \mathscr{B})$. Thus we may find $E \subseteq U$ with $U \in \sigma(\mathscr{A} \times \mathscr{B})$ with $(\mu \times \nu)(U) = 0$. By Lemma 5.8.1 we have

$$\int_X \phi_U \,\mathrm{d}\mu = 0.$$

Now let $Z = \{x \in X \mid \phi_U(x) \neq 0\}$. We must have $\mu(Z) = 0$. Thus, for almost every $x \in X$ we have $\mu(U_x) = 0$. Since $E_x \subseteq U_x$ and since μ is complete, it follows that E_x is $\mathscr{B}$-measurable for almost every $x \in X$. If $y \notin E_x$ then we must have $h_x(y) = 0$. This implies that, as long as $x \notin Z$ then h_x is measurable and zero almost everywhere. This completes the proof. ■

5.8.3 Fubini's Theorem

Now let us investigate swapping the order of integration in computing integrals on products. Let us see what we might mean by this. If $f\colon X \times Y \to \overline{\mathbb{R}}$ is $\sigma(\mathscr{A} \times \mathscr{B})$-measurable or $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$-measurable, then we define

$$\phi_f(x) = \begin{cases} \int_Y f_x \,\mathrm{d}\nu, & \text{the integral exists,} \\ 0, & \text{otherwise,} \end{cases} \qquad \psi_f(y) = \begin{cases} \int_X f^y \,\mathrm{d}\mu, & \text{the integral exists,} \\ 0, & \text{otherwise.} \end{cases}$$

We may then ask when it holds that

$$\int_X \phi_f \, d\mu = \int_Y \psi_f \, d\nu,$$

and when, if the preceding equality holds, both sides are, in fact, the integral of f with respect to the product measure. We have two more or less identical theorems, one for the product measure and one for its completion.

The first theorem deals with the product measure on $A \times B$.

5.8.4 Theorem (Fubini's Theorem for the product measure) *Let $(\mathsf{X}, \mathscr{A}, \mu)$ and $(\mathsf{Y}, \mathscr{B}, \nu)$ be σ-finite measure spaces and let $\mathsf{f}\colon \mathsf{A} \times \mathsf{B} \to \overline{\mathbb{R}}$ be $\sigma(\mathscr{A} \times \mathscr{B})$-measurable. Then the following statements hold:*

(i) *if f is $\overline{\mathbb{R}}_{\geq 0}$-valued then ϕ_f and ψ_f are measurable and*

$$\int_\mathsf{X} \phi_\mathsf{f} \, d\mu = \int_\mathsf{Y} \psi_\mathsf{f} \, d\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathsf{f} \, d(\mu \times \nu);$$

(ii) *if $\phi_{|\mathsf{f}|} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ or if $\psi_{|\mathsf{f}|} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu), \overline{\mathbb{R}})$, then*

$$\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X} \times \mathsf{Y}, \sigma(\mathscr{A} \times \mathscr{B}), \mu \times \nu); \overline{\mathbb{R}})$$

and

$$\int_\mathsf{X} \phi_\mathsf{f} \, d\mu = \int_\mathsf{Y} \psi_\mathsf{f} \, d\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathsf{f} \, d(\mu \times \nu);$$

(iii) *if $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X} \times \mathsf{Y}, \sigma(\mathscr{A} \times \mathscr{B}), \mu \times \nu); \overline{\mathbb{R}})$ then*

(a) *$\mathsf{f}_\mathsf{x} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \overline{\mathbb{R}})$ and $\mathsf{f}^\mathsf{y} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ for almost every $\mathsf{x} \in \mathsf{X}$ and $\mathsf{y} \in \mathsf{Y}$,*

(b) *$\phi_\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ and $\psi_\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \overline{\mathbb{R}})$, and*

(c) *it holds that*

$$\int_\mathsf{X} \phi_\mathsf{f} \, d\mu = \int_\mathsf{Y} \psi_\mathsf{f} \, d\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathsf{f} \, d(\mu \times \nu).$$

Proof (i) By Lemma 5.8.3 the functions ϕ_f and ψ_f are everywhere defined since the integral of a nonnegative-valued measurable function always exists. By Lemma 5.8.1 this part of the theorem holds for characteristic functions of $\mathscr{L}(A) \times \mathscr{L}(B)$-measurable sets. Therefore, it also holds for simple functions by Proposition 5.7.16 since simple functions are finite linear combinations of characteristic functions. By Proposition 5.6.39 let $(g_j)_{j\in\mathbb{Z}_{>0}}$ be a monotonically increasing sequence of simple functions such that $f(x, y) = \lim_{j\to\infty} g_j(x, y)$ for each $(x, y) \in X\times Y$. By the Monotone Convergence Theorem we have

$$\int_X \phi_f \, d\mu = \lim_{j\to\infty} \int_X \phi_{g_j} \, d\mu = \lim_{j\to\infty} \int_{X\times Y} g_j \, d(\mu \times \nu) = \int_{X\times Y} f \, d(\mu \times \nu),$$

and similarly for ψ_f. This gives the result.

(ii) By part (i) we have

$$\int_X \phi_{|f|}\,\mathrm{d}\mu = \int_X \psi_{|f|}\,\mathrm{d}\nu = \int_{X\times Y} |f|\,\mathrm{d}(\mu\times\nu) < \infty.$$

Thus f is $\mu\times\nu$-integrable, as desired. Note, then, that f_+, f_- are $\mu\times\nu$-integrable. Thus $f \in \mathsf{L}^{(1)}((X\times Y), \mathscr{A}\times\mathscr{B}, \mu\times\nu); \overline{\mathbb{R}})$ by Exercise 5.7.4. By part (i) we have

$$\int_X \phi_{f_+}\,\mathrm{d}\mu = \int_X \psi_{f_+}\,\mathrm{d}\nu = \int_{X\times Y} f_+\,\mathrm{d}(\mu\times\nu),$$

and similarly for f_-. By Proposition 5.7.17 it then follows that

$$\int_X \phi_f\,\mathrm{d}\mu = \int_X \psi_f\,\mathrm{d}\nu = \int_{X\times Y} f\,\mathrm{d}(\mu\times\nu),$$

as desired.

(iii) Write $f = f_+ - f_-$ and note that f_x, $f_{+,x}$, and $f_{-,x}$ are $\mathscr{B}$-measurable by Lemma 5.8.3. By part (i) the functions ϕ_{f_+} and ϕ_{f_-} are $\mathscr{A}$-measurable. Also by part (i) we have

$$\int_X \phi_{f_+}\,\mathrm{d}\mu = \int_X \psi_{f_+}\,\mathrm{d}\nu = \int_{X\times Y} f_+\,\mathrm{d}(\mu\times\nu),$$

and similarly for f_-. Therefore, ϕ_{f_+} and ϕ_{f_-} are integrable with respect to μ. Therefore, ϕ_{f_+} and ϕ_{f_-} are finite for almost all $x \in X$ by Proposition 5.7.12. If

$$Z = \{x \in X \mid \phi_{f_+}(x) = \infty\} \cup \{x \in X \mid \phi_{f_-}(x) = \infty\}$$

then $Z \in \mathscr{A}$ by Proposition 5.6.6 and $\mu(Z) = 0$. If $x \notin Z$ then we have

$$\phi_f(x) = \int_X f_+\,\mathrm{d}\mu - \int_X f_-\,\mathrm{d}\mu = \phi_{f_+}(x) - \phi_{f_-}(x)$$

and if $x \in Z$ we have $\phi_f(x) = 0$. Thus ϕ_f almost everywhere agrees with $\phi_{f_+} - \phi_{f_-}$. By Propositions 5.7.11 and 5.7.17 we have

$$\begin{aligned}\int_X \phi_f\,\mathrm{d}\mu &= \int_X \phi_{f_+}\,\mathrm{d}\mu - \int_X \phi_{f_-}\,\mathrm{d}\mu\\ &= \int_{X\times Y} f_+\,\mathrm{d}(\mu\times\nu) - \int_{X\times Y} f_-\,\mathrm{d}(\mu\times\nu)\\ &= \int_{X\times Y} f\,\mathrm{d}(\mu\times\nu),\end{aligned}$$

as desired. A similar argument gives

$$\int_Y \psi_f\,\mathrm{d}\nu = \int_{X\times Y} f\,\mathrm{d}(\mu\times\nu)$$

which completes the proof. ■

We shall also use the following result, which follows from the previous theorem, along with the definition of the integral for vector-valued functions. In the statement of the theorem, we use the obvious definitions for $\boldsymbol{f}_x$ and $\boldsymbol{f}^y$ for a function $\boldsymbol{f}\colon X\times Y \to \mathbb{R}^n$ and for functions $\phi_{\boldsymbol{f}}\colon X \to \mathbb{R}^n$ and $\psi_{\boldsymbol{f}}\colon Y \to \mathbb{R}^n$.

5.8.5 Corollary (Vector-valued Fubini's Theorem for the product measure) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *and* $(\mathsf{Y}, \mathscr{B}, \nu)$ *be σ-finite measure spaces and let* $\mathbf{f}\colon \mathsf{A} \times \mathsf{B} \to \mathbb{R}^{\mathrm{n}}$ *be* $\sigma(\mathscr{A} \times \mathscr{B})$*-measurable. Then the following statements hold:*

(i) *if* $\phi_{\|\mathbf{f}\|_{\mathbb{R}^{\mathrm{n}}}} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathrm{n}})$ *or if* $\psi_{\|\mathbf{f}\|_{\mathbb{R}^{\mathrm{n}}}} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu), \mathbb{R}^{\mathrm{n}})$*, then*

$$\mathbf{f} \in \mathsf{L}^{(1)}((\mathsf{X} \times \mathsf{Y}, \sigma(\mathscr{A} \times \mathscr{B}), \mu \times \nu); \mathbb{R}^{\mathrm{n}})$$

and

$$\int_{\mathsf{X}} \phi_{\mathbf{f}}\, \mathrm{d}\mu = \int_{\mathsf{Y}} \psi_{\mathbf{f}}\, \mathrm{d}\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathbf{f}\, \mathrm{d}(\mu \times \nu);$$

(ii) *if* $\mathbf{f} \in \mathsf{L}^{(1)}((\mathsf{X} \times \mathsf{Y}, \sigma(\mathscr{A} \times \mathscr{B}), \mu \times \nu); \mathbb{R}^{\mathrm{n}})$ *then*

(a) $\mathbf{f}_{\mathrm{x}} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \mathbb{R}^{\mathrm{n}})$ *and* $\mathbf{f}^{\mathrm{y}} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathrm{n}})$ *for almost every* $\mathrm{x} \in \mathsf{X}$ *and* $\mathrm{y} \in \mathsf{Y}$*,*

(b) $\phi_{\mathbf{f}} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \mathbb{R}^{\mathrm{n}})$ *and* $\psi_{\mathbf{f}} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \mathbb{R}^{\mathrm{n}})$*, and*

(c) *it holds that*

$$\int_{\mathsf{X}} \phi_{\mathbf{f}}\, \mathrm{d}\mu = \int_{\mathsf{Y}} \psi_{\mathbf{f}}\, \mathrm{d}\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathbf{f}\, \mathrm{d}(\mu \times \nu).$$

Of course, the theorem applies to the space case of $\mathbb{R}^2$ and so to $\mathbb{C}$-valued functions.

Let us give some examples that illustrate how to use Fubini's Theorem, as well as some of the caveats one must be aware of when applying the theorem.

5.8.6 Examples (Fubini's Theorem)

1. Let us take $X = Y = \mathbb{Z}_{>0}$, $\mathscr{A} = \mathscr{B} = \mathbf{2}^{\mathbb{Z}_{>0}}$, and $\mu = \nu = \mu_\Sigma$, where we recall from Example 5.3.9–3 that μ_Σ denotes the counting measure. For $f\colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{R}$ and $m \in \mathbb{Z}_{>0}$ define $f_m\colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{R}$ by

$$f_m(j,k) = \begin{cases} f(j,k), & j,k \in \{1,\dots,m\}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that

$$\int_{\mathbb{Z}_{>0}\times\mathbb{Z}_{>0}} f_m\, \mathrm{d}(\mu_\Sigma \times \mu_\Sigma) = \sum_{j=1}^{m}\sum_{k=1}^{m} f(j,k)$$

since f_m is a simple function. Clearly, $f(j,k) = \lim_{m\to\infty} f_m(j,k)$ for every $(j,k) \in \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$. Thus, by the Monotone Convergence Theorem,

$$\int_{\mathbb{Z}_{>0}\times\mathbb{Z}_{>0}} |f|\, \mathrm{d}\mu_\Sigma \times \mu_\Sigma = \lim_{m\to\infty} \sum_{j=1}^{m}\sum_{j=1}^{m} |f(j,k)|.$$

In other words, $f \in \mathsf{L}^{(1)}((\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}), \mathbf{2}^{\mathbb{Z}_{>0}} \times \mathbf{2}^{\mathbb{Z}_{>0}}, \mu_\Sigma \times \mu_\Sigma); \mathbb{R})$ if and only if

$$\sum_{j,k=1}^{\infty} |f(j,k)| < \infty,$$

noting that the doubly infinite sum is unambiguously defined since it is a sum of positive terms, cf. Theorem 2.4.5.

Now, Fubini's Theorem in this case tells us that when $f \in \mathsf{L}^{(1)}((\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}, \mathbf{2}^{\mathbb{Z}_{>0}} \times \mathbf{2}^{\mathbb{Z}_{>0}}, \mu_\Sigma \times \mu_\Sigma); \mathbb{R})$ then

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} f(j,k) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} f(j,k) = \sum_{j,k=1}^{\infty} f(j,k),$$

i.e., the order of summation can be swapped.

2. We take $X = Y = \mathbb{Z}$, $\mathscr{A} = \mathscr{B} = \mathbf{2}^{\mathbb{Z}}$, and $\mu = \nu = \mu_\Sigma$. We define $f \colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ by

$$f(j,k) = \begin{cases} 1, & j \in \mathbb{Z}_{\geq 0},\ k = j, \\ -1, & j \in \mathbb{Z}_{\geq 0},\ k = j+1, \\ 0, & \text{otherwise.} \end{cases}$$

We directly compute

$$\phi_f(j) = 0, \qquad \psi_f(k) = \begin{cases} 1, & k = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$\int_{\mathbb{Z}} \phi_f \, \mathrm{d}\mu_\Sigma = 0, \qquad \int_{\mathbb{Z}} \psi_f \, \mathrm{d}\mu_\Sigma = 1,$$

which shows that the order of integration (order of summation, in this case) cannot be swapped. This does not contradict Theorem 5.8.4, however. Indeed, note that

$$\phi_{|f|}(j) = \begin{cases} 2, & j \in \mathbb{Z}_{\geq 0}, \\ 0, & \text{otherwise,} \end{cases} \qquad \psi_{|f|} = \begin{cases} 1, & j = 0, \\ 2, & j \in \mathbb{Z}_{>0}, \\ 0, & \text{otherwise.} \end{cases}$$

Since neither of these functions is integrable, part (ii) of Theorem 5.8.4 cannot be applied.

3. One might wonder whether the fact that the measure spaces are infinite in the preceding example is the reason for the failure of Fubini's Theorem. In this example, we shall show that this is not the case. Here we shall use the Lebesgue integral, which is defined using the Lebesgue measure. Although we do not discuss this in detail until Sections 5.9 and **??**, this should not cause problems since for this example it suffices to consider the functions as being Riemann integrable.

We take $X = Y = [0,1]$, $\mathscr{A} = \mathscr{B} = \mathscr{L}([0,1])$, and $\mu = \nu = \lambda_{[0,1]}$. Define $\xi_j = 1 - \frac{1}{j+1}$, $j \in \mathbb{Z}_{>0}$, and let $g_j \colon [0,1] \to \mathbb{R}$ be a positive continuous function such that $\int_{[0,1]} g_j \, \mathrm{d}\lambda_{[0,1]} = 1$ and such that $\operatorname{supp}(g_j) \subseteq (\xi_j, \xi_{j+1})$ (for example, a "triangular" function of the right height and base). Then define $f \colon [0,1] \times [0,1] \to \mathbb{R}$ by

$$f(x,y) = \sum_{j=1}^{\infty} (g_j(x) - g_{j+1}(x)) g_j(y).$$

It is clear that for each $(x, y) \in [0,1] \times [0,1]$ this sum has at most one nonzero term, and so is well-defined. By construction, we have

$$\phi_f(x) = \sum_{j=1}^{\infty} (g_j(x) - g_{j+1}(x)) \int_{[0,1]} g_j \, \mathrm{d}\lambda_{[0,1]} = \sum_{j=1}^{\infty} (g_j(x) - g_{j+1}(x))$$

and

$$\psi_f(y) = \sum_{j=1}^{\infty} g_j(y) \int_{[0,1]} (g_j - g_{j+1}) \, \mathrm{d}\lambda_{[0,1]} = 0.$$

Therefore, observing that

$$\sum_{j=1}^{\infty} (g_j(x) - g_{j+1}(x)) = g_1(x),$$

we have

$$\int_{[0,1]} \phi_f \, \mathrm{d}\lambda_{[0,1]} = 1, \quad \int_{[0,1]} \psi_f \, \mathrm{d}\lambda_{[0,1]} = 0,$$

showing that the order of integration cannot be swapped. But this does not contradict part (ii) of Theorem 5.8.4 since

$$\begin{aligned} \phi_{|f|}(x) = \sum_{j=1}^{\infty} |g_j(x) - g_{j+1}(x)| \int_{[0,1]} g_j \, \mathrm{d}\lambda_{[0,1]} &= \sum_{j=1}^{\infty} (g_j(x) + g_{j+1}(x)) \\ \implies \quad \int_{[0,1]} \phi_{|f|} \, \mathrm{d}\lambda_{[0,1]} = \infty, \end{aligned}$$

using the fact that the functions g_j, $j \in \mathbb{Z}_{>0}$, are positive and have pairwise disjoint support. Thus the hypotheses of part (ii) of Theorem 5.8.4 do not hold.

4. Let us consider now a case where Fubini's Theorem can fail for a positive-valued function. Again, we make use of the Lebesgue integral. We take $X = Y = [0,1]$, $\mathscr{A} = \mathbf{2}^{[0,1]}$[a], $\mathscr{B} = \mathscr{L}([0,1])$, and $\mu = \mu_\Sigma$ and $\nu = \lambda_{[0,1]}$. We define $f \colon \mathbb{Z} \times \mathbb{R} \to \mathbb{R}$ by

$$f(x, y) = \begin{cases} 1, & x = y, \\ 0, & \text{otherwise.} \end{cases}$$

Then we compute

$$\phi_f(x) = 0, \quad \psi_f(y) = 1$$

for all $(x, y) \in [0,1] \times [0,1]$. Therefore,

$$\int_{[0,1]} \phi_f \, \mathrm{d}\mu_\Sigma = 0, \quad \int_{[0,1]} \psi_f \, \mathrm{d}\lambda_{[0,1]} = 1.$$

Again, the order of integration cannot be swapped. In this case, the issue cannot be with the hypotheses of part (ii) of Theorem 5.8.4 since f is nonnegative-valued, and so it is part (i) that should be applied. However, the problem with this example is that the measure space $([0,1], \mathbf{2}^{[0,1]}, \mu_\Sigma)$ is not σ-finite.

5. *missing stuff* •

Next we state the version of Fubini's Theorem for the completion of the product measure. This is actually the version of Fubini's Theorem that gets the most use since it applies to the Lebesgue integral on $\mathbb{R}^n$ as a product measure. Fortunately, it differs from Theorem 5.8.4 only in the use of the completed measure in the statement.

5.8.7 Theorem (Fubini's Theorem for the completion of the product measure) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *and* $(\mathsf{Y}, \mathscr{B}, \nu)$ *be σ-finite measure spaces and let* $\mathsf{f}\colon \mathsf{A} \times \mathsf{B} \to \overline{\mathbb{R}}$ *be* $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$*-measurable. Then the following statements hold:*

(i) if f *is* $\overline{\mathbb{R}}_{\geq 0}$*-valued then* ϕ_{f} *and* ψ_{f} *are measurable and*

$$\int_{\mathsf{X}} \phi_{\mathsf{f}}\, \mathrm{d}\mu = \int_{\mathsf{Y}} \psi_{\mathsf{f}}\, \mathrm{d}\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathsf{f}\, \mathrm{d}(\overline{\mu \times \nu});$$

(ii) if $\phi_{\mathsf{f}} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *or if* $\psi_{\mathsf{f}} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu), \overline{\mathbb{R}})$*, then*

$$\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X} \times \mathsf{Y}, \overline{\sigma}(\mathscr{A} \times \mathscr{B}), \overline{\mu \times \nu}); \overline{\mathbb{R}})$$

and

$$\int_{\mathsf{X}} \phi_{\mathsf{f}}\, \mathrm{d}\mu = \int_{\mathsf{Y}} \psi_{\mathsf{f}}\, \mathrm{d}\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathsf{f}\, \mathrm{d}(\overline{\mu \times \nu});$$

(iii) if $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X} \times \mathsf{Y}, \overline{\sigma}(\mathscr{A} \times \mathscr{B}), \overline{\mu \times \nu}); \overline{\mathbb{R}})$ *then*

(a) $\mathsf{f}_{\mathsf{x}} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \overline{\mathbb{R}})$ *and* $\mathsf{f}^{\mathsf{y}} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *for almost every* $\mathsf{x} \in \mathsf{X}$ *and* $\mathsf{y} \in \mathsf{Y}$*,*

(b) $\phi_{\mathsf{f}} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and* $\psi_{\mathsf{f}} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \overline{\mathbb{R}})$*, and*

(c) it holds that

$$\int_{\mathsf{X}} \phi_{\mathsf{f}}\, \mathrm{d}\mu = \int_{\mathsf{Y}} \psi_{\mathsf{f}}\, \mathrm{d}\nu = \int_{\mathsf{X}\times\mathsf{Y}} \mathsf{f}\, \mathrm{d}(\overline{\mu \times \nu}).$$

Proof By Proposition 5.7.15 we can write $f = g + h$ where

$$\mu \times \nu(\{(x, y) \in X \times Y \mid h(x, y) \neq 0\}) = 0$$

and where g is $\sigma(\mathscr{A} \times \mathscr{B})$-measurable. One now applies Theorem 5.8.4 to g, notes that $f_x = g_x$ for almost every x by Lemma 5.8.3, and therefore deduces from Proposition 5.7.15 that

$$\int_{X\times Y} f\, \mathrm{d}(\overline{\mu \times \nu}) = \int_X \phi_f\, \mathrm{d}\mu = \int_X \phi_g\, \mathrm{d}\mu = \int_{X\times Y} g\, \mathrm{d}(\mu \times \nu),$$

provided that all integrals exist. A similar conclusion holds using f^y, g^y, ψ_f, and ψ_g. The theorem follows directly from this. ■

The next result deals with a situation we will commonly encounter when using Fubini's theorem.

5.8.8 Corollary (A special case of Fubini's Theorem) *Let* $(\mathsf{X}, \mathscr{A}, \mu)$ *and* $(\mathsf{Y}, \mathscr{B}, \nu)$ *be* σ-*finite measure spaces, let* $\mathsf{f} \in \mathsf{L}^{(0)}((\mathsf{X}, \mathscr{A}); \overline{\mathbb{R}})$ *and* $\mathsf{g} \in \mathsf{L}^{(0)}((\mathsf{Y}, \mathscr{B}); \overline{\mathbb{R}})$*, and define* $\mathsf{F}\colon \mathsf{X} \times \mathsf{Y} \to \overline{\mathbb{R}}$ *by* $\mathsf{F}(\mathsf{x}, \mathsf{y}) = \mathsf{f}(\mathsf{x})\mathsf{g}(\mathsf{y})$*. Then*

(i) F *is both* $\sigma(\mathscr{A} \times \mathscr{B})$*- and* $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$*-measurable and*

(ii) F *is integrable with respect to both* $\mu \times \nu$ *and* $\overline{\mu \times \nu}$ *if* $\mathsf{f} \in \mathsf{L}^{(1)}((\mathsf{X}, \mathscr{A}, \mu); \overline{\mathbb{R}})$ *and* $\mathsf{g} \in \mathsf{L}^{(1)}((\mathsf{Y}, \mathscr{B}, \nu); \overline{\mathbb{R}})$.

Proof (i) Denote $\tilde{f}, \tilde{g}\colon X \times Y \to \mathbb{R}$ by $\tilde{f}(x, y) = f(x)$ and $\tilde{g}(x, y) = g(y)$. Then

$$\tilde{f}^{-1}([a, \infty]) = f^{-1}([a, \infty]) \times Y \in \mathscr{A} \times \mathscr{B},$$

and so both $\tilde{f}$ is $\sigma(\mathscr{A} \times \mathscr{B})$- and $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$-measurable. Similarly, $\tilde{g}$ is both $\sigma(\mathscr{A} \times \mathscr{B})$- and $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$-measurable. Therefore, by Proposition 5.6.11, $\tilde{f}\tilde{g}$ is both $\sigma(\mathscr{L}(A) \times \mathscr{L}(B))$- and $\overline{\sigma}(\mathscr{A} \times \mathscr{B})$-measurable. This part of the result follows since $F = \tilde{f}\tilde{g}$.

(ii) By part (i) of Theorem 5.8.4 we compute

$$\int_{X \times Y} |F| \, \mathrm{d}(\mu \times \nu) = \int_Y |g| \Big(\int_X |f| \, \mathrm{d}\mu \Big) \, \mathrm{d}\nu = \Big(\int_X |f| \, \mathrm{d}\mu \Big) \Big(\int_Y |g| \, \mathrm{d}\nu \Big) < \infty.$$

The result now follows from part (ii) of Theorem 5.8.4. ■

missing stuff

5.8.9 Example (Fubini's Theorem for the Lebesgue measure) Let us consider $X = Y = \mathbb{R}$, $\mathscr{A} = \mathscr{B} = \mathscr{L}(\mathbb{R})$, and $\mu = \nu = \lambda$. Define $f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$f(x, y) = \begin{cases} 1, & x \in \mathbb{R}_{\geq 0},\ y \in [x, x+1], \\ -1, & x \in \mathbb{R}_{>0},\ y \in [x+1, x+2], \\ 0, & \text{otherwise.} \end{cases}$$

In Sections 5.9 and **??** we shall show that the Lebesgue integral agrees with the Riemann integral in cases where the latter is defined. Therefore, to work out this example, it suffices to perform integration using the usual Riemann integral. Let us then denote the integral with respect to the first factor by $\int \mathrm{d}x$ and the integral with respect to the second factor by $\int \mathrm{d}y$. In Figure 5.5 we depict the function. With this figure in mind we compute

$$\begin{aligned} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x, y) \, \mathrm{d}x \Big) \mathrm{d}y &= \int_0^1 \Big(\int_0^y \mathrm{d}x \Big) \mathrm{d}y + \int_1^2 \Big(\int_{y-1}^y \mathrm{d}x - \int_0^{y-1} \Big) \mathrm{d}y \\ &\quad + \int_2^\infty \Big(\int_{y-1}^y \mathrm{d}x - \int_{y-2}^{y-1} \mathrm{d}x \Big) \mathrm{d}y \\ &= \frac{1}{2} + \frac{1}{2} + 0 = 1 \end{aligned}$$

and

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x, y) \, \mathrm{d}y \Big) \mathrm{d}x = \int_0^\infty \Big(\int_x^{x+1} \mathrm{d}y - \int_{x+1}^{x+2} \Big) \mathrm{d}x = 0.$$

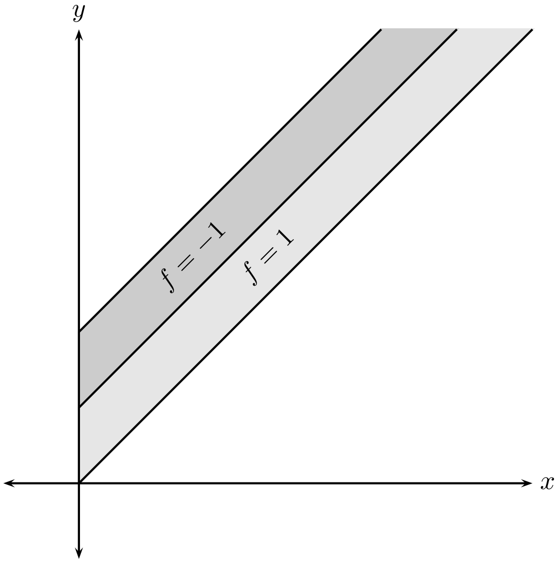


Figure 5.5 A function for which Fubini's Theorem does not hold

Thus both integrals

$$\int_{\mathbb{R}}\Big(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}x\Big)\,\mathrm{d}y, \qquad \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}y\Big)\,\mathrm{d}x$$

exist, but they are not equal to one another. However, this does not contradict Theorem 5.8.7. To see this, note that

$$\phi_{|f|}(x) = \begin{cases} 2, & x \in \mathbb{R}_{\geq 0}, \\ 0, & \text{otherwise}, \end{cases} \qquad \psi_{|f|}(y) = \begin{cases} y, & y \in [0,2], \\ y-2, & y \in (2,\infty), \\ 0, & \text{otherwise}. \end{cases}$$

Since neither of these functions is integrable, part (ii) of Theorem 5.8.7 does not apply. More directly, $|f|$ is the characteristic function of the union of the shaded regions in Figure 5.5. Therefore, the integral of $|f|$ is the area of this region which is infinity. Thus f is not integrable. •

Exercises

5.8.1

Section 5.9

The single-variable Lebesgue integral

The Lebesgue integral on $\mathbb{R}$ is nothing but the integral defined in Section 5.7.1 when the measure space is $(\mathbb{R}, \mathscr{L}(\mathbb{R}), \lambda)$. We shall not develop the *definition* of Lebesgue integral beyond this observation, so the reader looking to understand this definition will have to read Section 5.4 and then read Section 5.7 replacing all occurrences of $(X, \mathscr{A}, \mu)$ with $(\mathbb{R}, \mathscr{L}(\mathbb{R}), \lambda)$. This will give the reader most of what they will need to use the Lebesgue integral effectively. In this section we gather a few results and observations that are particuar to the Lebesgue integral on $\mathbb{R}$.

Do I need to read this section? The reader looking for the definition of the Lebesgue integral and some of its basic properties will get that by reading Sections 5.4 and 5.7 as described above. If this is all one is interested in, then this section can be bypassed, and the results consulted when needed. One topic in this section that may be of interest, and which is not contained in Sections 5.4 and 5.7, is the relationship between the Lebesgue integral and the Riemann integral. This, after all, is how we motivated the constructions that have gotten us to where we are. •

5.9.1 Lebesgue measurable functions

We begin by studying the character of Lebesgue measurable functions on $\mathbb{R}$. In this case, the additional structure of $\mathbb{R}$ allows us to give some further refinements of the properties of measurable functions.

Let us introduce the common terminology for the particular measurable functions we discuss in this section.

5.9.1 Definition (Borel measurable, Lebesgue measurable) Let $A \subseteq \mathbb{R}$. A function $f\colon A \to \overline{\mathbb{R}}$ is

(i) ***Borel measurable*** if $A \in \mathscr{B}(\mathbb{R})$ and if f is $\mathscr{B}(A)$-measurable and

(ii) ***Lebesgue measurable*** if $A \in \mathscr{L}(\mathbb{R})$ and if f is $\mathscr{L}(A)$-measurable.

We shall almost always write $\mathsf{L}^{(0)}(A; \overline{\mathbb{R}})$ for the Lebesgue measurable functions on A, rather than $\mathsf{L}^{(0)}((\mathbb{R}, \mathscr{L}(A)); \overline{\mathbb{R}})$. •

Now let us consider the approximation of measurable functions by "nice" functions like step functions and continuous functions. We recall from Section 3.4.1 the notion of a step function defined on a compact interval.

5.9.2 Theorem (Lebesgue measurable functions are approximated by step functions) *If* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *is a compact interval, if* $\mathrm{f}\colon \mathrm{I} \to \overline{\mathbb{R}}$ *is measurable and satisfies*

$$\lambda(\{\mathrm{x} \in \mathrm{I} \mid \mathrm{f}(\mathrm{x}) \in \{-\infty, \infty\}\} = 0,$$

and if $\epsilon_1, \epsilon_2 \in \mathbb{R}_{>0}$*, then there exists a step function* $\mathrm{g}\colon \mathrm{I} \to \mathbb{R}_{\geq 0}$ *such that*

$$\lambda(\{\mathrm{x} \in \mathrm{I} \mid |\mathrm{f}(\mathrm{x}) - \mathrm{g}(\mathrm{x})| \geq \epsilon_1\}) < \epsilon_2.$$

Proof It suffices to prove the theorem when $\epsilon_1 = \epsilon_2 = \epsilon$. Thus we take $\epsilon \in \mathbb{R}_{>0}$.

For $k \in \mathbb{Z}_{>0}$ define

$$A_k = \{x \in I \mid |f(x)| \geq k\},$$

and note that the sequence $(\lambda(I \setminus A_k))_{k\in\mathbb{Z}_{>0}}$ is monotonically increasing and bounded above by $b - a$. Thus it is convergent by Theorem 2.3.8. Moreover, it converges to $b - a$. Indeed, if the sequence converges to $\ell < b - a$ then this would imply, by Proposition 5.3.3, that

$$\lim_{k\to\infty} \lambda(I \setminus A_k) = \lambda(I \setminus \cup_{k\in\mathbb{Z}_{>0}} A_k) < b - a.$$

Thus there exists a set $B \subseteq I$ of positive measure such that $I = (\cup_{k\in\mathbb{Z}_{>0}A_k} \overset{\circ}{\cup} B)$. Note if $x \in B$ then $|f(x)| = \infty$, contradicting our assumptions on f. Thus we indeed have $\lim_{k\to\infty} \lambda(I \setminus A_k) = b - a$. Thus there exists $M \in \mathbb{Z}_{>0}$ such that $\lambda(I \setminus A_M) < b - a - \frac{\epsilon}{2}$, i.e., $\lambda(A_M) < \frac{\epsilon}{2}$. Therefore,

$$\lambda(\{x \in I \mid |f(x)| \geq M\}) < \tfrac{\epsilon}{2}.$$

Then define $f_M\colon I \to \mathbb{R}$ by

$$f_M(x) = \begin{cases} f(x), & |f(x)| < M, \\ M, & |f(x)| \geq M, \\ -M, & f(x) < -M. \end{cases}$$

Note that f_M is measurable by Proposition 5.6.16.

Now take $K \in \mathbb{Z}_{>0}$ such that $2^{-K} < \epsilon$ and such that $K \geq M$. If we follow the construction in the proof of Proposition 5.6.39 then we define

$$A_{+,K,j} = \{x \in I \mid 2^{-K}(j-1) \leq f_M(x) < 2^{-K}j\}$$

and

$$A_{-,K,j} = \{x \in I \mid -2^{-K}j \leq f_M(x) < -2^{-K}(j-1)\}$$

for $j \in \{1, \dots, K2^K\}$. Since $K \geq M$ we have

$$I = (\cup_{j=1}^{K2^K} A_{+,K,j}) \cup (\cup_{j=1}^{K2^K} A_{-,K,j}).$$

Moreover, if we define a simple function $h\colon I \to \mathbb{R}$ by

$$h(x) = \begin{cases} 2^{-K}(j-1), & x \in A_{+,K,j}, \\ -2^{-K}j, & x \in A_{-,K,j}, \end{cases}$$

then we have $|h(x) - f_M(x)| < \epsilon$ for every $x \in I$.

Now that we have a $\mathbb{R}$-valued simple function h that approximates f_M to within ϵ on I, let us dispense with the cumbersome notation above we introduced to define h, and instead write $h = \sum_{j=1}^{k} a_j \chi_{A_j}$ for $a_1, \dots, a_k \in \mathbb{R}$ and for a partition $(A_1, \dots, A_k)$

of I into Lebesgue measurable sets. Fix $j \in \{1, \dots, k\}$. Since A_j is measurable, by Corollary 5.4.20 we can write $A_j = U_j \setminus B_j$ where U_j is open and where $B_j \subseteq U_j$ satisfies $\lambda(B_j) < \frac{\epsilon}{8k}$. Since U_j is open, it is a countable union of disjoint open intervals by Proposition 2.5.6. If U_j is in fact a finite union of open intervals then denote $V_j = U_j$. If any of the intervals comprising V_j have common endpoints, then these intervals may be shrunk so that their complement in A_j has measure at most $\frac{\epsilon}{2k}$. Next suppose that U_j is a countable union of open intervals $(J_{j,l})_{l \in \mathbb{Z}_{>0}}$. Since U_j is bounded we must have $\sum_{l=1}^{\infty} \lambda(J_{j,l}) < \infty$. Therefore, there exists $N_j \in \mathbb{Z}_{>0}$ such that $\sum_{j=N_j+1}^{\infty} \lambda(J_{l,j}) < \frac{\epsilon}{8k}$. We then define $V_j = \cup_{l=1}^{N_j} J_{j,l}$. If any of the intervals $J_{1,j}, \dots, J_{N_j+1,j}$ have common endpoints, they can be shrunk while maintaining the fact that the measure of their complement in A_j is at most $\frac{\epsilon}{2k}$. Define $g \colon I \to \mathbb{R}$ on V_j by asking that $g(x) = a_j$ for $x \in V_j$. Doing this for each $j \in \{1, \dots, k\}$ defines $g \colon I \to \mathbb{R}$ on the set $\cup_{j=1}^{k} V_j$ which is a finite union of open intervals whose complement has measure at most $\frac{\epsilon}{2}$. The complement to $\cup_{j=1}^{k} V_j$ is a union of intervals, and on these intervals define g to be, say, 0. Note that g as constructed is a step function, and that $g(x) = h(x)$ for $x \in \cup_{j=1}^{k} V_j$.

Note that if $x \in (\cup_{j=1}^{k} V_j) \cup (I \setminus A_M)$ we have

$$|g(x) - f(x)| = |h(x) - f_M(x)| < \epsilon.$$

Therefore,

$$\lambda(\{x \in I \mid f(x) - g(x) \geq \epsilon\}) \subseteq I \setminus ((\cup_{j=1}^{k} V_j) \cup (I \setminus A_M)),$$

and

$$\lambda(I \setminus ((\cup_{j=1}^{k} V_j) \cup (I \setminus A_M))) < \epsilon,$$

giving the result. ■

A similar sort of result holds for approximations of measurable functions by continuous functions.

5.9.3 Theorem (Lebesgue measurable functions are approximated by continuous functions) *If* $I = [a, b]$ *is a compact interval, if* $f \colon I \to \overline{\mathbb{R}}$ *is measurable and satisfies*

$$\lambda(\{x \in I \mid f(x) \in \{-\infty, \infty\}\} = 0,$$

and if $\epsilon_1, \epsilon_2 \in \mathbb{R}_{>0}$*, then there exists a continuous function* $h \colon I \to \mathbb{R}_{\geq 0}$ *such that*

$$\lambda(\{x \in I \mid |f(x) - h(x)| \geq \epsilon_1\}) < \epsilon_2.$$

Proof We shall merely outline how this works, since this is "obvious" once one has the basic idea at hand. We assume that $\epsilon_1 = \epsilon_2 = \epsilon$. By the method of Theorem 5.9.2, we approximate f with a step function g such that

$$\lambda(\{x \in I \mid |f(x) - g(x)| \geq \epsilon\}) < \epsilon.$$

Note that the set of points in I where $|f(x) - g(x)| < \epsilon$ is a finite union of intervals with pairwise disjoint closures on each of which g is constant. The value of g on the intervals complementary to these intervals is of no consequence. To define the continuous function h we ask that h agree with g on the intervals upon which g is constant, and between these intervals we ask that h be a linear function that interpolates between the values of h at the two endpoints. The resulting function clearly satisfies the conclusions of the theorem. ■

In Definition 6.7.28 we will***missing stuff*** define the support for continuous functions as the closure of the set of points where the function is nonzero. For continuous functions, this is a satisfactory definition. For more general classes of functions, this is not so. For example, if $f\colon \mathbb{R} \to \mathbb{R}$ is the characteristic function of $\mathbb{Q}$, then the definition of support for continuous functions, when applied to f, gives $\operatorname{supp}(f) = \mathbb{R}$. However, this does not reflect the fact that f is zero almost everywhere. So we adapt the notion of support for continuous functions to measurable functions as follows.

5.9.4 Definition (Support of a measurable function) Let $f \in \mathsf{L}^{(0)}(\mathbb{R}; \overline{\mathbb{R}})$ and define

$$\mathscr{O}_f = \{U \subseteq \mathbb{R} \mid U \text{ open and } f(x) = 0 \text{ for almost every } x \in U\}.$$

Then the ***support*** of f is $\operatorname{supp}(f) = \mathbb{R} \setminus (\cup_{U \in \mathscr{O}_f} U)$. •

Being the complement of an open set, the support of a measurable function is closed. The following result gives the essential property of closure.

5.9.5 Proposition (Characterisation of support) *For* $f, g \in \mathsf{L}^{(0)}(\mathbb{R}; \overline{\mathbb{R}})$, *the following two statements hold:*

(i) $f(x) = 0$ *for almost every* $x \in \mathbb{R} \setminus \operatorname{supp}(f)$*;*

(ii) if $f(x) = g(x)$ *for almost every* $x \in \mathbb{R}$ *then* $\operatorname{supp}(f) = \operatorname{supp}(g)$.

Proof (i) We have $\mathbb{R} \setminus \operatorname{supp}(f) = \mathscr{O}_f$ in the notation of Definition 5.9.4. Recall from Definition **??** that the distance between $x \in \mathbb{R}$ and $A \subseteq \mathbb{R}$ is denoted by

$$\operatorname{dist}(x, A) = \inf\{|y - x| \mid y \in A\}.$$

Let $k \in \mathbb{Z}_{>0}$ and define

$$K_k = \{x \in \mathscr{O}_f \mid \operatorname{dist}(x, \operatorname{supp}(f)) \geq \tfrac{1}{k},\ |x| \leq k\}.$$

By Proposition **??**, the function $x \mapsto \operatorname{dist}(x, A)$ is continuous. By Corollary 3.1.4, since the set $[\frac{1}{k}, \infty)$ is closed and since $\overline{\mathsf{B}}(k, 0)$ is closed, K_k is the intersection of closed sets, and so closed. Therefore, since it is also bounded, it is compact. Since $K_k \subseteq \mathscr{O}_f$ and since $\mathscr{O}_f$ is a union of open sets, by the Heine–Borel Theorem, K_k is a finite union of open sets from $\mathscr{O}_f$, say $K_k = \cup_{j=1}^{m_k} U_{k,j}$. Denote

$$Z_{k,j} = \{x \in U_{k,j} \mid f(x) \neq 0\}, \qquad j \in \{1, \dots, m_k\}.$$

Since $f(x) = 0$ for almost every $x \in U_{j_k}$ for each $j_k \in \{1, \dots, m_k\}$, it follows that $\lambda(Z_{k,j}) = 0$. Therefore, since the set of points in K_k at which f is nonzero is $\cup_{j=1}^{m_k} Z_{k,j}$, it follows that $f(x) = 0$ for almost every $x \in K_k$. Now note that $\mathscr{O}_f = \cup_{k \in \mathbb{Z}_{>0}} K_k$. Thus the set of points $x \in \mathscr{O}_f$ such that $f(x) \neq 0$ is a countable union of sets of measure zero, and so has measure zero. That is, $f(x) = 0$ for almost every $x \in \mathscr{O}_f$.

(ii) We claim that if f and g agree almost everywhere, then $\mathscr{O}_f = \mathscr{O}_g$. Indeed, suppose that $U \in \mathscr{O}_f$ so that $f(x) = 0$ for almost every $x \in U$. Define

$$Z_1 = \{x \in U \mid f(x) \neq 0\}, \qquad Z_2 = \{x \in U \mid g(x) \neq f(x)\}.$$

Note that Z_1 and Z_2 have measure zero and so $Z_1 \cup Z_2$ also has measure zero. Moreover, if $x \in U \setminus (Z_1 \cup Z_2)$ then $g(x) = f(x) = 0$. Thus $U \in \mathscr{O}_g$ and so $\mathscr{O}_f \subseteq \mathscr{O}_g$. Reversing the argument shows that $\mathscr{O}_g \subseteq \mathscr{O}_f$. It then immediately follows that $\operatorname{supp}(f) = \operatorname{supp}(g)$. ■

Let us give an example which shows that the notion of support must be treated with some care, the previous result notwithstanding.

5.9.6 Example (A caveat concerning the support of a function) Note that $\mathbb{Q} \subseteq \mathbb{R}$ has Lebesgue measure zero. It follows, by definition of measure zero, that there exists a countable collection of intervals $((a_j, b_j))_{j\in\mathbb{Z}_{>0}}$ such that

$$\sum_{j=1}^{\infty} |b_j - a_j| < 1$$

and such that $\mathbb{Q} \subseteq \cup_{j\in\mathbb{Z}_{>0}}(a_j, b_j)$. Let us define $A = \cup_{j\in\mathbb{Z}_{>0}}(a_j, b_j)$. By countable subadditivity of the Lebesgue measure we have $\lambda(A) \leq 1$. We claim that $\mathrm{supp}(\chi_A) = \mathbb{R}$. Indeed, if $U \in \mathscr{O}_{\chi_A}$ then $\lambda(A \cap U) = o$. If U is nonempty then it contains an interval, say (a, b). Note that A is a nonempty open set by Exercise 2.5.1. Moreover, since there are rational numbers in (a, b) by Proposition 2.2.15, it follows that $A \cap U$ is a nonempty open set, and so has positive Lebesgue measure. We conclude, therefore, that if $U \in \mathscr{O}_{\chi_A}$ then $U = \emptyset$. Thus $\mathrm{supp}(\chi_A) = \mathbb{R}$, as claimed. The point is that we have

$$\lambda(A) \leq 1 < \lambda(\mathrm{supp}(A)) = \infty.$$

Thus the measure of the support of a function can far exceed the measure of the set of points where the function is nonzero. This is a consequence of our asking that the support be a closed set. •

For continuous functions, the preceding definition of support reduces to the usual one, i.e., the one used in Definition 6.7.28.

5.9.7 Proposition (The support of a continuous function) *If* $\mathrm{f}\colon \mathbb{R} \to \mathbb{R}$ *is continuous then*

$$\mathrm{supp(f) = cl(\{x \in \mathbb{R} \mid f(x) \neq 0\})}.$$

Proof Let $x_0 \in \mathbb{R} \setminus \mathrm{supp}(f)$. Then there exists $U \in \mathscr{O}_f$ such that $x_0 \in U$. By Exercise 3.1.12 we have $f(x) = 0$ for every $x \in U$. In particular, $f(x_0) = 0$ and, moreover, $f(x) = 0$ in the neighbourhood U of x_0. Thus x_0 cannot be a limit $\lim_{j\to\infty} x_j$ with $f(x_j) \neq 0$. That is,

$$x_0 \notin \mathrm{cl}(\{x \in \mathbb{R} \mid f(x) \neq 0\}).$$

Conversely, suppose that $x_0 \in \mathbb{R} \setminus \mathrm{cl}(\{x \in \mathbb{R} \mid f(x) \neq 0\})$. Then there must be a neighbourhood U of x_0 such that $f(x) = 0$ for every $x \in U$. Thus $U \subseteq \mathscr{O}_f$ and so $x \in \mathbb{R} \setminus \mathrm{supp}(f)$. ■

5.9.2 The (conditional) Lebesgue integral

Let $\mathscr{L}(\mathbb{R})$ be the collection of Lebesgue measurable subsets of $\mathbb{R}$ (see Definition 5.4.4) and let $\lambda\colon \mathscr{L}(\mathbb{R}) \to \overline{\mathbb{R}}_{\geq 0}$ be the Lebesgue measure (see Definition 5.4.4). From Proposition 5.4.6, recall also that if $A \in \mathscr{L}(\mathbb{R})$ then we denote by $\mathscr{L}(A)$ the Lebesgue measurable subsets of A and by λ_A the restriction of λ to $\mathscr{L}(A)$.

Although it is pretty clear if you have been reading this chapter from the beginning, perhaps the following definition ought to be made for those who "skipped to the good bit."

5.9.8 Definition (Lebesgue integral on $\mathbb{R}$) If $f \in \mathsf{L}^{(0)}(\mathbb{R}; \overline{\mathbb{R}})$ then f is ***Lebesgue integrable*** and the ***Lebesgue integral*** of f is the integral of f with respect to the Lebesgue measure when the integral exists:

$$\int_{\mathbb{R}} f \, \mathrm{d}\lambda.$$

If $f \in \mathsf{L}^{(0)}(A; \overline{\mathbb{R}})$, then f is ***Lebesgue integrable*** and the ***Lebesgue integral*** of f is the integral of f with respect to the Lebesgue measure when the integral exists:

$$\int_{A} f \, \mathrm{d}\lambda_A.$$

We shall almost always denote the Lebesgue integrable functions on A by $\mathsf{L}^{(1)}(A; \overline{\mathbb{R}})$ rather than $\mathsf{L}^{(1)}((A, \mathscr{L}(A), \lambda_A); \overline{\mathbb{R}})$. •

Of course, if $A \in \mathscr{L}(\mathbb{R})$ and if $f \in \mathsf{L}^{(0)}(A; \overline{\mathbb{R}})$, we can think of f as being in $\mathsf{L}^{(0)}(\mathbb{R}; \overline{\mathbb{R}})$ by making it zero outside A. The resulting function can be directly verified to be measurable (cf. Exercise 5.6.3). We can, therefore, write

$$\int_{A} f \, \mathrm{d}\lambda_A = \int_{\mathbb{R}} f \, \mathrm{d}\lambda$$

without risk of confusion. When it is convenient to do so, we shall do this. We will also omit the subscript "A" in "$\mathrm{d}\lambda_A$" when the resulting compactness of notation is desired. Thus, we will use the symbols

$$\int_{A} f \, \mathrm{d}\lambda_A, \quad \int_{A} f \, \mathrm{d}\lambda, \quad \int_{\mathbb{R}} f \, \mathrm{d}\lambda$$

to stand for the same thing when it is clear from context what is meant.

It is worth making some connections at this point with how we defined the single-variable Riemann integral in Section 3.4. For the Riemann integral we had two constructions which we showed were equivalent when the domain of the function was a compact interval. However, the so-called conditional Riemann integral generalises the Riemann integral when the domain of the function is a not a compact interval. This can be generalised for the Lebesgue integral as follows.

5.9.9 Definition (Conditionally Lebesgue integrable functions on a general interval) Let $I \subseteq \mathbb{R}$ be an interval and let $f\colon I \to \overline{\mathbb{R}}$ be a function whose restriction to every compact subinterval of I is Lebesgue integrable.

(i) If $I = [a, b]$ then define

$$C\int_{I} f \, \mathrm{d}\lambda_I = \int_{I} f \, \mathrm{d}\lambda.$$

(ii) If $I = (a, b]$ then define

$$C\int_{I} f \, \mathrm{d}\lambda_I = \lim_{r_a \downarrow a} \int_{[r_a, b]} f \, \mathrm{d}\lambda_{[r_a, b]}$$

if the limit exists.

(iii) If $I = [a, b)$ then define

$$C\int_I f\,\mathrm{d}\lambda_I = \lim_{r_b \uparrow b} \int_{[a,r_b]} f\,\mathrm{d}\lambda_{[a,r_b]}$$

if the limit exists.

(iv) If $I = (a, b)$ then define

$$C\int_I f\,\mathrm{d}\lambda_I = \lim_{r_a \downarrow a} \int_{[r_a,c]} f\,\mathrm{d}\lambda_{[r_a,c]} + \lim_{r_b \uparrow b} \int_{[c,r_b]} f\,\mathrm{d}\lambda_{[c,r_b]}$$

for some $c \in (a, b)$, if the limit exists.

(v) If $I = (-\infty, b]$ then define

$$C\int_I f\,\mathrm{d}\lambda_I = \lim_{R \to \infty} \int_{[-R,b]} f\,\mathrm{d}\lambda_{[-R,b]}$$

if the limit exists.

(vi) If $I = (-\infty, b)$ then define

$$C\int_I f\,\mathrm{d}\lambda_I = \lim_{R \to \infty} \int_{[-R,c]} f\,\mathrm{d}\lambda_{[-R,c]} + \lim_{r_b \uparrow b} \int_{[c,r_b]} f\,\mathrm{d}\lambda_{[c,r_b]}$$

for some $c \in (-\infty, b)$, if the limit exists.

(vii) If $I = [a, \infty)$ then define

$$C\int_I f\,\mathrm{d}\lambda_I = \lim_{R \to \infty} \int_{[a,R]} f\,\mathrm{d}\lambda_{[a,R]}$$

if the limit exists.

(viii) If $I = (a, \infty)$ then define

$$C\int_I f\,\mathrm{d}\lambda_I = \lim_{r_a \downarrow a} \int_{[r_a,c]} f\,\mathrm{d}\lambda_{[r_a,c]} + \lim_{R \to \infty} \int_{[c,R]} f\,\mathrm{d}\lambda_{[c,R]}$$

for some $c \in (a, \infty)$, if the limit exists.

(ix) If $I = \mathbb{R}$ then define

$$C\int_{\mathbb{R}} f\,\mathrm{d}\lambda = \lim_{R \to \infty} \int_{[-R,c]} f\,\mathrm{d}\lambda_{[-R,c]} + \lim_{R \to \infty} \int_{[c,R]} f\,\mathrm{d}\lambda_{[c,R]}$$

for some $c \in \mathbb{R}$, if the limit exists.

If, for a given I and f, the appropriate of the above limits exists, then f is ***conditionally Lebesgue integrable*** on I, and the ***conditional Lebesgue integral*** is the value of the limit. •

It is not usual to define the conditional Lebesgue integral, but we do so in order to make our analogies with the Riemann integral, explored in Section 5.9.3, more clear. Thus a few comments are relevant at this point.

5.9.10 Remarks (On the conditional Lebesgue integral)

1. Since the Lebesgue integral is so general, it is not really natural to restrict the definition of the Lebesgue integral to functions defined on intervals. Indeed, a somewhat more natural construction would be as follows. Let $A \in \mathscr{L}(\mathbb{R})$ and let $f\colon A \to \overline{\mathbb{R}}$ be measurable. By Theorem 5.4.19 let $(K_j)_{j\in\mathbb{Z}_{>0}}$ be a family of compact sets such that $K_j \subseteq A$, $K_j \subseteq K_{j+1}$, $j \in \mathbb{Z}_{>0}$, and $\lambda(A) = \lim_{j\to\infty} \lambda(K_j)$. Then we can define the conditional Lebesgue integral of f by
$$C\int_A f\,\mathrm{d}\lambda_A = \lim_{j\to\infty} \int_{K_j} (f|K_j)\,\mathrm{d}\lambda_{K_j}.$$
This construction generalises the more complicated, but more direct construction of Definition 5.9.9. Since we will not use this level of generality for the conditional Lebesgue integral, we shall stick to the more concrete Definition 5.9.9 as our definition of the conditional Lebesgue integral. It also make more clear the comparison with the Riemann integral.
2. The conditional Lebesgue integral shares with the Lebesgue integral the usual properties with respect to operations on functions, i.e., those properties given in Section 5.7.2 for the general integral. The verification of this is a matter of using the results of Section 5.7.2, the fact that the conditional Lebesgue integral is defined as a limit, and the fact that limits commute with natural operations as shown in Section 2.3.6. We leave the details of proving this statement to a sufficiently bored reader. However, we shall make free use of these facts ourselves.
3. In Theorem 5.9.11 we shall show that the (conditional) Lebesgue integral generalises the (conditional) Riemann integral. For this reason, to give an example of a function that is conditionally Lebesgue integrable but not Lebesgue integrable, it suffices to give an example of a function that is conditionally Riemann integrable but not Riemann integrable. Such a function is given in Example 3.4.20.

•

5.9.3 Properties of the Lebesgue integral

In this section we shall give some useful properties of the Lebesgue integral and the conditional Lebesgue integral. In the preceding section we constructed two versions of the Lebesgue integral for functions of a single variable. As was pointed out in the course of these constructions, these two integral mirror in spirit the development in Section 3.4 for the Riemann integral. We begin this section by showing that the Riemann integral is generalised by the Lebesgue integral.

One of the intentions of Section 5.1 was to show that the Riemann integral suffers a few theoretical defects. If the Lebesgue integral is to redress these problems, it would be helpful it applied in all cases when the Riemann integral applies. This is indeed the case.

5.9.11 Theorem (The (conditional) Lebesgue integral generalises the (conditional) Riemann integral) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is (conditionally) Riemann* *integrable, then* f *is (conditionally) Lebesgue integrable, and*

$$(\mathrm{C})\int_{\mathrm{I}} \mathrm{f(x)\,dx} = (\mathrm{C})\int_{\mathrm{I}} \mathrm{f}\,\mathrm{d}\lambda.$$

Proof First let us consider the case where $I = [a,b]$ is compact. Suppose that $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. For $k \in \mathbb{Z}_{>0}$ let P_k be a partition with the property that $A_+(f,P_k) - A_-(f,P_k) < \frac{1}{k}$. By redefining partitions if necessary we can assume that the endpoints of the intervals for P_{k+1} contain those for P_k, cf. Lemma 1 from the proof of Theorem 3.4.9. Upon doing this, the sequences $(s_+(f,P_k)(x))_{k\in\mathbb{Z}_{>0}}$ and $(s_-(f,P_k)(x))_{k\in\mathbb{Z}_{>0}}$ are increasing and decreasing, respectively, for each $x \in [a,b]$. Moreover, since the functions in these sequences are step functions, they are simple functions and so are measurable. It is also clear that the Riemann integral of a step function is equal to the Lebesgue integral of the same function, by definition of the Riemann integral of a step function and the Lebesgue integral of a simple function. Thus

$$\int_{[a,b]} s_+(f,P_k)\,\mathrm{d}\lambda = A_+(f,P_k), \qquad \int_{[a,b]} s_-(f,P_k)\,\mathrm{d}\lambda = A_-(f,P_k).$$

Denote

$$f_+(x) = \lim_{k\to\infty} s_+(f,P_k)(x), \quad f_-(x) = \lim_{k\to\infty} s_-(f,P_k)(x),$$

for $x \in [a,b]$. Proposition 5.6.18 implies that f_+ and f_- are measurable. Note that f, and therefore f_+ and f_-, are bounded. Thus f_+ and f_- are bounded in absolute value by a constant function. Such a function is obviously in $\mathscr{L}^{(1)}([a,b];\mathbb{R})$, and so the Dominated Convergence Theorem implies that

$$\int_{[a,b]} f_+\,\mathrm{d}\lambda = \lim_{k\to\infty} A_+(f,P_k) = \int_a^b f(x)\,\mathrm{d}x$$

and

$$\int_{[a,b]} f_-\,\mathrm{d}\lambda = \lim_{k\to\infty} A_-(f,P_k) = \int_a^b f(x)\,\mathrm{d}x,$$

where we have used the characterisation of the Riemann integral in Theorem 3.4.9. From this we conclude that

$$\int_{[a,b]} (f_+ - f_-)\,\mathrm{d}\lambda = 0,$$

which implies that $f_+(x) = f_-(x)$ for almost every $x \in [a,b]$ by Proposition 5.7.14. Since $f_-(x) \le f(x) \le f_+(x)$ for every $x \in []1,b]$, it, therefore, follows that f is itself measurable (being almost everywhere equal to the measurable functions f_+ and f_-) and Lebesgue integrable (again, being almost everywhere equal to the Lebesgue integrable functions f_+ and f_-). Moreover, by Proposition 5.7.11 it follows that

$$\int_{[a,b]} f_+\,\mathrm{d}\lambda = \int_{[a,b]} f_-\,\mathrm{d}\lambda = \int_{[a,b]} f\,\mathrm{d}\lambda = \int_a^b f(x)\,\mathrm{d}x,$$

as desired.

Now we consider an arbitrary interval $I \subseteq \mathbb{R}$ and suppose that f is Riemann integrable. Here, we first take f to be nonnegative-valued. In this case, the definition of the Riemann integral from Definition 3.4.14 implies that there exists a sequence $(I_k)_{k\in\mathbb{Z}_{>0}}$ of compact intervals such that $I_k \subseteq I_{k+1}$, $k \in \mathbb{Z}_{>0}$, such that $I = \cup_{k\in\mathbb{Z}_{>0}} I_k$, and such that

$$\int_I f(x)\,\mathrm{d}x = \lim_{k\to\infty} \int_{I_k} f(x)\,\mathrm{d}x.$$

From the Monotone Convergence Theorem, Theorem 5.7.24, and the first part of the proof it then follows that

$$\int_I f\,\mathrm{d}\lambda = \lim_{k\to\infty} \int_{I_k} f(x)\,\mathrm{d}x = \int_I f(x)\,\mathrm{d}x.$$

For general $\mathbb{R}$-valued f, the result follows from writing $f = f_+ - f_-$, and using linearity of the Riemann and Lebesgue integrals, Propositions 3.4.22 and 5.7.17.

Finally, we consider an arbitrary interval I and suppose that f is conditionally Riemann integrable. According to Definition 5.9.9 there exists a sequence $(K_j = [a_j, b_j])_{j\in\mathbb{Z}_{>0}}$ of compact intervals such that $K_j \subseteq K_{j+1}$, $j \in \mathbb{Z}_{>0}$, and such that $I \cup_{j\in\mathbb{Z}_{>0}} K_j$. By our arguments above we have

$$\int_{K_j} (f|K_j)\,\mathrm{d}\lambda_{K_j} = \int_{a_j}^{b_j} f(x)\,\mathrm{d}x, \qquad j \in \mathbb{Z}_{>0}.$$

Therefore,

$$\lim_{j\to\infty} \int_{K_j} (f|K_j)\,\mathrm{d}\lambda_{K_j} = \lim_{j\to\infty} \int_{a_j}^{b_j} f(x)\,\mathrm{d}x,$$

and the result follows by the definitions of the conditional Riemann and Lebesgue integrals. ■

We must, of course, also show that there are Lebesgue integrable functions that are not Riemann integrable.

5.9.12 Example (A Lebesgue integrable, but not Riemann integrable, function) Let $I = [0,1]$ and let $A = \mathbb{Q} \cap [0,1]$. Then define $f\colon [0,1] \to \mathbb{R}$ by $f = \chi_A$. Note that f is not Riemann integrable; see Example 3.4.10. However, f is Lebesgue integrable, as can be seen in many ways. Most directly, f is the characteristic function of the Lebesgue measurable set A, and so is Lebesgue integrable simply by definition. If one wishes, one can also "derive" the Lebesgue integrability of f. For example, if we let $(q_k)_{k\in\mathbb{Z}_{>0}}$ be an enumeration of the set A, we can define $g_k\colon [0,1] \to \mathbb{R}$ by

$$g_k(x) = \begin{cases} 1, & x \in \{q_1, \dots, q_k\}, \\ 0, & \text{otherwise.} \end{cases}$$

The functions g_k, $k \in \mathbb{Z}_{>0}$, are Lebesgue integrable, indeed Riemann integrable, cf. Example 5.1.11. Moreover, $f(x) = \lim_{k\to\infty} g_k(x)$ for all $x \in [0,1]$. By the Dominated Convergence Theorem (verify its hypotheses!), we then have

$$\int_{[0,1]} f\,\mathrm{d}\lambda = \lim_{k\to\infty} \int_{[0,1]} g_k\,\mathrm{d}\lambda = \lim_{k\to\infty} \int_0^1 g_k(x)\,\mathrm{d}x = 0.$$

Thus f is indeed Lebesgue integrable, with Lebesgue integral zero. •

It is rather important not to overstate the importance of this example. It is not interesting, but it does serve to easily verify that the Lebesgue integral generalises the Riemann integral.

5.9.13 Notation and Remarks (Riemann integral versus Lebesgue integral) Having now established the relationship between the Riemann and Lebesgue integrals, we shall often use the sometimes more convenient notation for the Riemann integral when we actually are using the Lebesgue integral. Thus, for example, we may well write

$$\int_a^b f(x)\,\mathrm{d}x, \quad \int_{-\infty}^b f(x)\,\mathrm{d}x, \quad \mathrm{C}\int_a^\infty f(x)\,\mathrm{d}x$$

where we really mean

$$\int_{[a,b]} f\,\mathrm{d}\lambda_{[a,b]}, \quad \int_{(-\infty,b]} f\,\mathrm{d}\lambda_{(-\infty,b]}, \quad C\int_{[a,\infty)} f\,\mathrm{d}\lambda_{[a,\infty)},$$

respectively.

This confounding of notation for the Lebesgue and Riemann integrals suggests that the additional generality of the Lebesgue integral is not of great importance. This is both true and not true. It *is* true that we shall not encounter specific examples of Lebesgue integrable functions that are not Riemann integrable. That is to say, we shall not often care to compute the Lebesgue integral in cases where the Riemann integral will not suffice. However, it *is* the case that the Riemann integral has certain undesirable features, as we discussed in Section 5.1.2. These undesirable features come in two basic flavours.

1. The Riemann and Lebesgue integrals both possess a Dominated Convergence Theorem, Theorems **??** and 5.7.28, respectively. However, the two theorems differ in a crucial way. Specifically, in the Dominated Convergence Theorem for the Riemann integral, the Riemann integrability of the limit function is an hypothesis, while in the Dominated Convergence Theorem for the Lebesgue integral, the integrability of the limit function is a conclusion. This inability of the Dominated Convergence Theorem for the Riemann integral to predict the integrability of the limit function is a crucial defect. We shall discuss this further in Section 5.9.11.
2. It is interesting to consider not just individual Riemann or Lebesgue integrable functions, but the *set* of all Riemann or Lebesgue integrable functions. We have already denoted by $\mathsf{L}^{(1)}(I;\mathbb{R})$ the set of $\mathbb{R}$-valued Lebesgue integrable functions on the interval I. Let us denote by $\mathsf{R}^{(1)}(I;\mathbb{R})$ the set of $\mathbb{R}$-valued Riemann integrable functions on I, cf. the discussion preceding Proposition 5.1.12. Both $\mathsf{L}^{(1)}(I;\mathbb{R})$ and $\mathsf{R}^{(1)}(I;\mathbb{R})$ are $\mathbb{R}$-vector spaces by the standard linearity properties of the integral. In Chapter 6 we shall discuss the notion of a normed vector space and the important related notion of completeness. We shall show in Theorem 6.7.56 (essentially) that the set of Lebesgue integrable functions form a complete normed vector space. This is not the case for Riemann integrable functions, as we show in Proposition 5.1.12. It may not be clear at this point

why this is important, but this is, in fact, *extremely* important. As we go along, and we use the Lebesgue integral at various points in these volumes, we shall point out instances where the particular properties of the Lebesgue integral are crucial. •

Now that we have established the close relationship between the Lebesgue and Riemann integrals, let us explore some of the properties of Lebesgue integrable functions. In Section 5.9.1 we explored the manner in which Lebesgue measurable functions can be pointwise approximated by "nice" functions like step functions or continuous functions. Lebesgue integrable functions, being Lebesgue measurable, are subject to the same approximations. However, for Lebesgue integrable functions we have another sort of approximation that is possible by virtue of the integral.

5.9.14 Theorem (Lebesgue integrable functions are approximated by step functions) *If* $\mathrm{I} = [\mathrm{a}, \mathrm{b}]$ *is a compact interval, if* $\mathrm{f} \in \mathsf{L}^{(1)}(\mathrm{I}; \overline{\mathbb{R}})$*, and if* $\epsilon \in \mathbb{R}_{>0}$*, then there exists a step function* $\mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *such that*

$$\int_{\mathrm{I}} |\mathrm{f} - \mathrm{g}| \, \mathrm{d}\lambda_{\mathrm{I}} < \epsilon.$$

Proof Let us first consider the case when f is bounded. Let $M \in \mathbb{R}_{>0}$ be such that $f(x) \leq M$ for all $x \in I$. Let $\epsilon \in \mathbb{R}_{>0}$. By Theorem 5.9.2 there exists a continuous function $g\colon I \to \mathbb{R}_{\geq 0}$ such that

$$\lambda\Big(\Big\{x \in I \;\Big|\; |f(x) - g(x)| < \tfrac{\epsilon}{(2(b-a))}\Big\}\Big) < \frac{\epsilon}{2M}.$$

Then

$$\int_I |f(x) - g(x)| \, \mathrm{d}\lambda_I < \frac{\epsilon}{2(b-a)}(b-a) + \frac{\epsilon}{2M} M < \epsilon,$$

giving the result in this case.

Next we consider the case when f is possibly unbounded and takes values in $\overline{\mathbb{R}}_{\geq 0}$. Let $\epsilon \in \mathbb{R}_{>0}$. For $M \in \mathbb{R}_{>0}$ define

$$f_M(x) = \begin{cases} f(x), & f(x) \leq M, \\ M, & f(x) > M. \end{cases}$$

Since $f \in \mathsf{L}^{(1)}(I; \overline{\mathbb{R}})$ we have $f(x) = \lim_{M\to\infty} f_M(x)$ for almost every $x \in I$. By the Dominated Convergence Theorem,

$$\lim_{M\to\infty} \int_I (f - f_M) \, \mathrm{d}\lambda_I = 0.$$

Thus there exists M sufficiently large that

$$\int_I |f(x) - f_M(x)| \, \mathrm{d}\lambda_I < \frac{\epsilon}{2}.$$

By the argument in the previous paragraph there exists a step function $g\colon I \to \mathbb{R}_{\geq 0}$ such that

$$\int_I |f_M - g| \, d\lambda_I < \frac{\epsilon}{2}.$$

Then, using the triangle inequality and monotonicity of the integral, Proposition 5.7.19,

$$\int_I |f - g| \, d\lambda_I \leq \int_I |f - f_M| \, d\lambda_I + \int_I |f_M - g| \, d\lambda_I < \epsilon,$$

giving the result in this case.

Finally, if f is $\overline{\mathbb{R}}$-valued, we write $f = f_+ - f_-$ for f_+ and f_- taking values in $\overline{\mathbb{R}}_{\geq 0}$. Let $\epsilon \in \mathbb{R}_{>0}$. By our arguments above there exists step functions $g_+, g_-\colon I \to \mathbb{R}_{\geq 0}$ such that

$$\int_I |f_+ - g_+| \, d\lambda_I < \frac{\epsilon}{2}, \quad \int_I |f_- - g_-| \, d\lambda_I < \frac{\epsilon}{2}.$$

Taking $g = g_+ - g_-$, the triangle inequality and Proposition 5.7.19 then give

$$\int_I |f - g| \, d\lambda_I \leq \int_I |f_+ - g_+| \, d\lambda_I + \int_I |f_- - g_-| \, d\lambda_I < \epsilon,$$

as desired. ■

A similar result as the previous holds for approximations of integrable functions by continuous functions. However, in this case it is possible to even be more general in terms of the domain of definition of the functions involved. The notion of support is used in the title of this theorem, but will only be introduced in Definition 6.7.28.

5.9.15 Theorem (Lebesgue integrable functions are approximated by compactly supported continuous functions) *If* $I \subseteq \mathbb{R}$ *is an interval, if* $f \in \mathsf{L}^{(1)}(I; \overline{\mathbb{R}})$*, and if* $\epsilon \in \mathbb{R}_{>0}$*, then there exists a continuous function* $g\colon I \to \mathbb{R}$ *such that*

$$\int_I |f - g| \, d\lambda_I < \epsilon$$

and such that the support of f*, i.e., the set*

$$\mathrm{cl}_I(\{x \in I \mid f(x) \neq 0\}),$$

is compact.

Proof If I is compact, then the result follows just like Theorem 5.9.14, but using Theorem 5.9.3 rather than Theorem 5.9.2. Thus the result holds when I is compact.

Thus we need only consider the case when I is not compact. Let $\epsilon \in \mathbb{R}_{>0}$. We let $(I_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence of compact intervals such that $I_j \subseteq I_{j+1}$ for each $j \in \mathbb{Z}_{>0}$ and such that $\cup_{j \in \mathbb{Z}_{>0}} I_j = I$. Define a sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathsf{L}^{(1)}(I; \mathbb{R})$ by

$$f_j(x) = \begin{cases} f(x), & x \in I_j, \\ 0, & \text{otherwise.} \end{cases}$$

By the Monotone Convergence Theorem we have

$$\lim_{j\to\infty}\int_I |f-f_j|\,\mathrm{d}\lambda_I=\int_I \lim_{j\to\infty}|f-f_j|\,\mathrm{d}\lambda_I=0.$$

Thus $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to f in $\mathsf{L}^{(1)}(I;\mathbb{F})$. Now, for each $j\in\mathbb{Z}_{>0}$, the fact that the theorem holds for compact intervals ensures the existence of a continuous function $h_j\colon I_j\to\mathbb{R}_{\geq 0}$ such that

$$\int_{I_j}|f_j|I_j-h_j|\,\mathrm{d}\lambda_{I_j}<\frac{\epsilon}{4}.$$

Note that if we extend h_j to I by asking that it be zero on $I\setminus I_j$ then this extension may not be continuous. However, we can linearly taper h_j to zero on $I\setminus I_j$ to arrive at a continuous function $g_j\colon I\to\mathbb{R}_{\geq 0}$ with compact support satisfying

$$\int_{I\setminus I_j}|g_j|\,\mathrm{d}\lambda_{I\setminus I_j}<\frac{\epsilon}{4}.$$

Then

$$\int_I|f_j-g_j|\,\mathrm{d}\lambda_I=\int_{I_j}|f_j-h_j|\,\mathrm{d}\lambda_{I_j}+\int_{I\setminus I_j}|g_j(x)|\,\mathrm{d}\lambda_{I\setminus I_j}<\frac{\epsilon}{4}+\frac{\epsilon}{4}<\frac{\epsilon}{2}.$$

Now choose $N\in\mathbb{Z}_{>0}$ sufficiently large that

$$\int_I|f-f_j|\,\mathrm{d}\lambda_I<\frac{\epsilon}{2}.$$

Then, by the triangle inequality,

$$\int_I|f-g_j|\,\mathrm{d}\lambda_I\leq\int_I|f-f_j|\,\mathrm{d}\lambda_I+\int_I|f_j-g_j|\,\mathrm{d}\lambda_I<\epsilon,$$

as desired. ■

5.9.4 Swapping operations with the Lebesgue integral

It is useful to have at hand results that tell us the nature of an integral as a function of a parameter. Thus we let $A\in\mathscr{L}(\mathbb{R})$ and let (a,b) be an open interval. We suppose that $f\colon(a,b)\times A\to\mathbb{R}$ has the property that, for $p\in(a,b)$, the function $x\mapsto f(p,x)$ is integrable. We denote $f^p(x)=f(p,x)$ and $f_x(p)=f(p,x)$. We then define

$$I_f(p)=\int_A f^p(x)\,\mathrm{d}x.$$

The next result indicates when such a function is continuous or differentiable.

5.9.16 Theorem (Continuous and differentiable dependence of integral on a parameter) *Let* $(\mathrm{a},\mathrm{b})\subseteq\mathbb{R}$*, let* $\mathrm{A}\in\mathscr{L}(\mathbb{R})$*, and let* $\mathrm{f}\colon(\mathrm{a},\mathrm{b})\times\mathrm{A}\to\mathbb{R}$ *have the property that* $\mathrm{f}^{\mathrm{p}}\in\mathsf{L}^{(1)}(\mathrm{A};\mathbb{R}))$ *for every* $\mathrm{p}\in(\mathrm{a},\mathrm{b})$*. Let* $\mathrm{p}_0\in(\mathrm{a},\mathrm{b})$*.*

(i) If f_{x} *is continuous at* p_0 *for almost every* $\mathrm{x}\in\mathrm{A}$ *and if there exists* $\mathrm{g}\in\mathsf{L}^{(1)}(\mathrm{A};\mathbb{R})$ *and a neighbourhood* U *of* p_0 *in* (a,b) *for which* $|\mathrm{f}^{\mathrm{p}}(\mathrm{x})|\leq\mathrm{g}(\mathrm{x})$ *for all* $\mathrm{p}\in\mathrm{U}$*, then* I_{f} *is continuous at* p_0*.*

(ii) *If there exists* $\epsilon \in \mathbb{R}_{>0}$ *so that*

(a) $(\mathrm{p}_0 - \epsilon, \mathrm{p}_0 + \epsilon) \subseteq (\mathrm{a}, \mathrm{b})$*,*

(b) f^{p} *is differentiable on* $(\mathrm{p}_0 - \epsilon, \mathrm{p}_0 + \epsilon)$*, and*

(c) *there exists* $\mathrm{g} \in \mathsf{L}^{(1)}(\mathrm{A}; \mathbb{R})$ *so that* $\left|\frac{\partial \mathrm{f}}{\partial \mathrm{p}}(\mathrm{p}, \mathrm{x})\right| \leq \mathrm{g}(\mathrm{x})$ *for* $\mathrm{p} \in (\mathrm{p}_0 - \epsilon, \mathrm{p}_0 + \epsilon)$ *and for almost every* $\mathrm{x} \in \mathrm{A}$*,*

then I_{f} *is differentiable at* p_0 *and*

$$\mathrm{I}'_{\mathrm{f}}(\mathrm{p}_0) = \int_{\mathrm{A}} \frac{\partial \mathrm{f}}{\partial \mathrm{p}}(\mathrm{p}_0, \mathrm{x})\, \mathrm{d}\mathrm{x}.$$

Proof (i) Let $(p_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in U this neighbourhood converging to p_0. By the Dominated Convergence Theorem we have

$$\lim_{j \to \infty} \int_A f(p_j, x)\, \mathrm{d}x = \int_A \lim_{j \to \infty} f(p_j, x)\, \mathrm{d}x = \int_A f(p_0, x)\, \mathrm{d}x,$$

the final equality by continuity of f_x for almost every $x \in A$ and by Theorem 3.1.3. This shows that $\lim_{j \to \infty} I_f(p_j) = I_f(p_0)$, giving the result by another application of Theorem 3.1.3.

(ii) We again let $(p_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence approaching p_0. By the Mean Value Theorem, for each $j \in \mathbb{Z}_{>0}$, there exists q_j between p_j and p_0 such that

$$\frac{f(p_j, x) - f(p_0, x)}{p_j - p_0} = \frac{\partial f}{\partial p}(q_j, x).$$

Note that we necessarily have $\lim_{j \to \infty} q_j = p_0$. Then we compute

$$\begin{aligned}\lim_{j \to \infty} \frac{I_f(p_j) - I_f(p_0)}{p_j - p_0} &= \lim_{j \to \infty} \int_A \frac{f(p_j, x) - f(p_0, x)}{p_j - p_0}\, \mathrm{d}x = \int_A \lim_{j \to \infty} \frac{f(p_j, x) - f(p_0, x)}{p_j - p_0}\, \mathrm{d}x \\ &= \int_A \lim_{j \to \infty} \frac{\partial f}{\partial p}(q_j, x)\, \mathrm{d}x = \int_A \frac{\partial f}{\partial p}(p_0, x)\, \mathrm{d}x.\end{aligned}$$

Here the interchanging of the limit and the integral is valid by the Dominated Convergence Theorem. ■

The above theorem is proved using tools that we presently have at our disposal. It suffices for many purposes. However, it is possible to weaken the hypotheses significantly while retaining the same conclusions, but at a price of using the notion of absolute continuity we introduce in Section 5.9.6 and the formalism of distributions we introduce in Chapter 10.

5.9.17 Theorem (A strong theorem on differential dependence of integral on a parameter) *Let* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *and let* $\mathrm{f}\colon \mathbb{R} \times \mathrm{A} \to \mathbb{R}$ *have the properties*

(i) *that* f_{x} *is locally absolutely continuous for almost every* $\mathrm{x} \in \mathrm{A}$ *and*

(ii) *that, for every compact subset* $\mathrm{K} \subseteq \mathbb{R}$*, the functions*

$$(\mathrm{p}, \mathrm{x}) \mapsto \mathrm{f}(\mathrm{p}, \mathrm{x}), \quad (\mathrm{p}, \mathrm{x}) \mapsto \mathbf{D}_1\mathrm{f}(\mathrm{p}, \mathrm{x}),$$

when restricted to $\mathrm{K} \times \mathrm{A}$*, are integrable.*

Then, if $\mathrm{I_f}\colon \mathbb{R} \to \mathbb{R}$ *is as above,* $\mathrm{I_f}$ *is locally absolutely continuous and*

$$\mathrm{I}'_{\mathrm{f}}(\mathrm{p}) = \int_{\mathrm{A}} \boldsymbol{D}_1 \mathrm{f}(\mathrm{p}, \mathrm{x})\,\mathrm{d}\mathrm{x}$$

for almost every $\mathrm{p} \in \mathrm{I}$.

Proof For $x \in A$ let $\theta_f(x) \in \mathscr{D}'(\mathbb{R};\mathbb{R})$ be the regular distribution associated with f_x. Adopting and slightly modifying the notation used in Proposition 10.2.43, let us define $F_f\colon A \times \mathscr{D}(\mathbb{R};\mathbb{R}) \to \mathbb{R}$ by

$$F_f(x,\phi) = \langle \theta_f(x); \phi \rangle = \int_{\mathbb{R}} f(p,x)\phi(p)\,\mathrm{d}p,$$

for $\phi \in \mathscr{D}(\mathbb{R};\mathbb{R})$ define $F_{f,\phi}\colon A \to \mathbb{R}$ by

$$F_{f,\phi}(x) = F_f(x,\phi),$$

and then define $\Theta_f\colon \mathscr{D}(\mathbb{R};\mathbb{R}) \to \mathbb{R}$ by

$$\Theta_f(\phi) = \int_A F_{f,\phi}\,\mathrm{d}\lambda = \int_A \Big(\int_{\mathbb{R}} f(p,x)\phi(p)\,\mathrm{d}p \Big)\,\mathrm{d}x.$$

We first claim that $\Theta_f \in \mathscr{D}'(\mathbb{R};\mathbb{R})$. We prove this by verifying that the hypotheses of Proposition 10.2.43 are satisfied. Let $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence converging to zero in $\mathscr{D}(\mathbb{R};\mathbb{R})$. Let $K \subseteq \mathbb{R}$ be a compact interval for which $\mathrm{supp}(\phi_j) \subseteq K$ for every $j \in \mathbb{Z}_{>0}$. Let

$$M = \sup\{|\phi_j(p)| \mid j \in \mathbb{Z}_{>0},\ p \in \mathbb{R}\},$$

noting that $M < \infty$ since the sequence $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ converges uniformly to zero. Then we have

$$\begin{aligned}\int_A |\sup\{F_{f,\phi_j}(x) \mid j \in \mathbb{Z}_{>0}\}|\,\mathrm{d}x &\le \int_A \Big(\int_{\mathbb{R}} \sup\{|f(p,x)\phi_j(p)| \mid j \in \mathbb{Z}_{>0}\}\,\mathrm{d}p \Big)\,\mathrm{d}x \\ &\le M \int_A \Big(\int_K |f(p,x)|\,\mathrm{d}p \Big)\,\mathrm{d}x < \infty,\end{aligned}$$

by hypothesis.

Now, for $\phi \in \mathscr{D}(\mathbb{R};\mathbb{R})$ we compute

$$\begin{aligned}\Theta'_f(\phi) &= -\Theta_f(\phi') = -\int_A \Big(\int_{\mathbb{R}} f(p,x)\phi'(p)\,\mathrm{d}p \Big)\,\mathrm{d}x \\ &= \int_A \Big(\int_{\mathbb{R}} \boldsymbol{D}_1 f(p,x)\phi(p)\,\mathrm{d}p \Big)\,\mathrm{d}x = \int_{\mathbb{R}} \Big(\int_A \boldsymbol{D}_1 f(p,x)\,\mathrm{d}x \Big)\phi(p)\,\mathrm{d}p\end{aligned}$$

using Proposition 5.9.34, the fact that ϕ has compact support, and Fubini's Theorem. This shows that Θ'_f is equal to the regular distribution associated with the function

$$p \mapsto \int_A \boldsymbol{D}_1 f(p,x)\,\mathrm{d}x.$$

By Proposition 10.2.31 it follows that this function is locally absolutely continuous and that it is equal almost everywhere to the derivative of the function

$$p \mapsto \int_A f(p,x)\,\mathrm{d}x,$$

which is the desired result. ■

It is also useful to have at hand a result which indicates when holomorphicity of the integrand implies holomorphicity of the integral.

5.9.18 Theorem (Holomorphic dependence on a parameter) *Suppose we have the following data:*

(i) a measurable subset $\mathrm{A} \subseteq \mathbb{R}$*;*

(ii) an open subset $\mathrm{D} \subseteq \mathbb{C}$*;*

(iii) a function $\mathrm{G}\colon \mathrm{A} \times \mathrm{D} \to \mathbb{C}$ *such that*

(a) the function $\mathrm{x} \mapsto \mathrm{G}(\mathrm{x}, \mathrm{z})$ *is in* $\mathsf{L}^{(1)}(\mathrm{A}; \mathbb{C})$ *for each* $\mathrm{z} \in \mathrm{D}$*,*

(b) the function $\mathrm{z} \mapsto \mathrm{G}(\mathrm{x}, \mathrm{z})$ *is in* $\mathsf{H}(\mathrm{D}, \mathbb{C})$ *for each* $\mathrm{x} \in \mathrm{A}$*, and*

(c) for each $\mathrm{z}_0 \in \mathrm{D}$ *there exists a neighbourhood* U *of* z_0 *in* D *and* $\mathrm{h} \in \mathsf{L}^{(1)}(\mathrm{A}; \mathbb{R}_{\geq 0})$ *such that* $|\mathrm{G}(\mathrm{x}, \mathrm{z})| \leq \mathrm{h}(\mathrm{x})$ *for each* $\mathrm{z} \in \mathrm{U}$*.*

Then the function $\mathrm{F}\colon \mathrm{D} \to \mathbb{C}$ *defined by*

$$\mathrm{F}(\mathrm{z}) = \int_{\mathrm{A}} \mathrm{G}(\mathrm{x}, \mathrm{z})\, \mathrm{d}\mathrm{x}$$

is in $\mathsf{H}(\mathrm{D}; \mathbb{C})$*.*

Proof By Theorem 5.9.16 we know that F is continuous in D. Now let Γ be a closed contour in D. Parameterise Γ with a map $\gamma\colon [0, L] \to D$***missing stuff*** so that

$$\int_{\Gamma} F(z)\, \mathrm{d}z = \int_0^L \Big(\int_A G(x, \gamma(s))\, \mathrm{d}x\Big)\, \mathrm{d}s.$$

Then the function $x \mapsto G(x, \gamma(s))$ is in $\mathsf{L}^{(1)}(A; \mathbb{C})$ for every $s \in [0, L]$. Also, the function $s \mapsto G(x, \gamma(s))$ is in $\mathsf{L}^{(1)}([0, L]; \mathbb{C})$ for every $x \in A$ since it is a continuous function defined on a compact interval. Therefore, Fubini's Theorem gives

$$\int_{\Gamma} F(z)\, \mathrm{d}z = \int_A \Big(\int_0^L G(x, \gamma(s))\, \mathrm{d}s\Big)\, \mathrm{d}x = \int_A \Big(\int_{\Gamma} G(x, z)\, \mathrm{d}z\Big) = 0,$$

using Cauchy's Theorem and holomorphicity of $z \mapsto G(x, z)$. Since this holds for every closed contour in D, Morera's Theorem allows us to conclude that F is holomorphic in D.***missing stuff*** ■

5.9.5 Locally Lebesgue integrable functions

Very often on wants to speak of functions that are integrable about every point, but which may not be integrable on their entire domain. This is another instance of the concept of "locality" that we have encountered many times before.

5.9.19 Definition (Locally Lebesgue integrable function) If $A \in \mathscr{L}(\mathbb{R})$ then $f \in \mathsf{L}^{(0)}(A; \overline{\mathbb{R}})$ is ***locally Lebesgue integrable***, or merely ***locally integrable***, if, for every compact set $K \subseteq A$, $f|K \in \mathsf{L}^{(1)}(K; \overline{\mathbb{R}})$. The set of locally Lebesgue integrable functions on A is denoted by $\mathsf{L}^{(1)}_{\mathrm{loc}}(A; \overline{\mathbb{R}})$. •

Note that if $f \in \mathsf{L}^{(0)}(A; \overline{\mathbb{R}})$ then one can define $\bar{f}\colon \mathbb{R} \to \overline{\mathbb{R}}$ to be defined to be equal to f on A and zero elsewhere. Moreover, $f \in \mathsf{L}^{(1)}_{\mathrm{loc}}(A; \overline{\mathbb{R}})$ if and only if $\bar{f} \in \mathsf{L}^{(1)}_{\mathrm{loc}}(\mathbb{R}; \overline{\mathbb{R}})$. Therefore, when talking about locally integrable functions one can, without loss of generality, think about functions whose domain is $\mathbb{R}$. When it is convenient to do this, we shall.

It is obvious that if f is integrable then it is locally integrable. Let us give some examples which clarify the meaning of local integrability as opposed to integrability.

5.9.20 Examples (Local integrability)

1. The function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is locally integrable (its restriction to every compact set is continuous and bounded) but not integrable.
2. The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^{-1/2}, & x \in \mathbb{R}_{>0}, \\ 0, & \text{otherwise}, \end{cases}$$

   is locally integrable but not integrable.
3. The function $f\colon [0, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^{-1/2}, & x \in (0, 1], \\ 0, & x = 0, \end{cases}$$

   is both locally integrable and integrable.
4. The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^{-1}, & x \in (0, 1], \\ 0, & \text{otherwise}, \end{cases}$$

   is both locally integrable and integrable. •

The following characterisation of locally integrable functions is sometimes useful.

5.9.21 Proposition (Characterisation of locally Lebesgue integrable functions) *For a function* $\mathrm{f}\colon \mathbb{R} \to \overline{\mathbb{R}}$ *the following statements are equivalent:*

(i) f *is locally Lebesgue integrable;*

(ii) for each $\mathrm{x} \in \mathbb{R}$ *there exists a neighbourhood* U *of* x *such that* $\mathrm{f}|\mathrm{U} \in \mathsf{L}^{(1)}(\mathrm{U}; \overline{\mathbb{R}})$*;*

(iii) for every continuous function $\mathrm{g}\colon \mathbb{R} \to \mathbb{R}$ *such that* $\mathrm{supp}(\mathrm{g})$ *is compact, it holds that* $\mathrm{fg} \in \mathsf{L}^{(1)}(\mathbb{R}; \overline{\mathbb{R}})$.

Proof (i) $\Longrightarrow$ (ii) Let $x \in \mathbb{R}$ and let $K = [x-1, x+1]$. By hypothesis, $f|K \in \mathsf{L}^{(1)}(K; \overline{\mathbb{R}})$ and so $f \in \mathsf{L}^{(1)}((x-1, x+1); \overline{\mathbb{R}})$ by Proposition 5.7.22. This gives the result with $U = (x-1, x+1)$.

(ii) $\Longrightarrow$ (iii) Let $g\colon \mathbb{R} \to \mathbb{R}$ be continuous with compact support. For $x \in \operatorname{supp}(g)$ there exists a neighbourhood U_x of x such that $f \in \mathsf{L}^{(1)}(U_x, \overline{\mathbb{R}})$. Since $(U_x)_{x \in \operatorname{supp}(g)}$ covers the compact set $\operatorname{supp}(g)$ there exists $x_1, \dots, x_k \in \operatorname{supp}(g)$ such that $\operatorname{supp}(g) \subseteq \cup_{j=1}^{k} U_{x_j}$. Since g is continuous with compact support there exists $M \in \mathbb{R}_{>0}$ such that $|g(x)| \leq M$ for every $x \in \mathbb{R}$ by Theorem 3.1.22. Then

$$\int_{\mathbb{R}} |fg| \, \mathrm{d}\lambda \leq M \int_{\operatorname{supp}(g)} f \, \mathrm{d}\lambda_{\operatorname{supp}(g)} \leq M \sum_{j=1}^{k} \int_{U_{x_j}} f \, \mathrm{d}\lambda_{U_{x_j}} < \infty,$$

since $\operatorname{supp}(g) \subseteq \cup_{j=1}^{k} U_{x_j}$. This gives the desired conclusion.

(iii) $\Longrightarrow$ (i) Let $K \subseteq \mathbb{R}$ be compact and let $a, b \in \mathbb{R}$ be such that $K \subseteq [a, b]$. Now take $g\colon \mathbb{R} \to \mathbb{R}$ defined by

$$g(x) = \begin{cases} 1, & x \in [a, b], \\ x - (a-1), & x \in [a-1, a), \\ -x + (b+1), & x \in (b, b+1], \\ 0, & \text{otherwise}. \end{cases}$$

Note that g is positive, continuous with compact support, and $g(x) = 1$ for all $x \in [a, b]$. Then

$$\int_{K} |f| \, \mathrm{d}\lambda_K \leq \int_{[a,b]} |f| \, \mathrm{d}\lambda_{[a,b]} \leq \int_{\mathbb{R}} |fg| \, \mathrm{d}\lambda < \infty,$$

giving the result. ■

Using the preceding characterisation of locally integrable functions, one can easily prove that the set of locally integrable functions is a subspace of the measurable functions.

5.9.22 Proposition (Algebraic operations on locally integrable functions) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$, *if* $\mathrm{f}, \mathrm{g} \in \mathsf{L}^{(1)}_{\mathrm{loc}}(\mathrm{A}; \overline{\mathbb{R}})$, *and if* $\mathrm{a} \in \mathbb{R}$, *then*

(i) $\mathrm{f} + \mathrm{g} \in \mathsf{L}^{(1)}_{\mathrm{loc}}(\mathrm{A}; \overline{\mathbb{R}})$ *and*

(ii) $\mathrm{af} \in \mathsf{L}^{(1)}_{\mathrm{loc}}(\mathrm{A}; \overline{\mathbb{R}})$.

Proof This follows from the definition of local integrability, along with Proposition 5.7.17. ■

Local integrability is not preserved by products and quotients, cf. Example 5.7.18.

5.9.6 Absolute continuity

In this section we introduce a special class of continuous functions that are almost everywhere differentiable. With this class of functions one can prove a stronger form of the Fundamental Theorem of Calculus than was possible when we initially discussed this in Section 3.4.6.

The definition of absolute continuity shares with the definition of bounded variation the feature of being unbearably cryptic at first sight. However, we shall see as we go along that absolute continuity is a notion that arises naturally from the Lebesgue integral.

5.9.23 Definition ((Locally) absolutely continuous function) Let $[a,b]$ be a compact interval. A function $f\colon [a,b] \to \mathbb{R}$ is ***absolutely continuous*** if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $\delta \in \mathbb{R}_{>0}$ such that, if $((a_j,b_j))_{j\in\{1,\dots,k\}}$ is a finite family of disjoint open intervals for which

$$\sum_{j=1}^{k} |b_j - a_j| < \delta,$$

then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < \epsilon.$$

For a general interval $I \subseteq \mathbb{R}$, a function $f\colon I \to \mathbb{R}$ is ***locally absolutely continuous*** if $f|J$ is absolutely continuous for every compact interval $J \subseteq I$. •

We can make the same sort of comments concerning "absolute continuity" versus "local absolute continuity" as were made in Notation 3.3.8 concerning the relationship between "bounded variation" and "locally bounded variation."

The following result gives the most basic properties of absolutely functions.

5.9.24 Proposition (Locally absolutely continuous functions are continuous and of locally bounded variation) *If* I $\subseteq \mathbb{R}$ *is an interval and if* f: I $\to \mathbb{R}$ *is a locally absolutely continuous function, then* f *is continuous and has locally bounded variation.*

Proof We first consider the case where $I = [a,b]$. Let $x \in [a,b]$ and let $\epsilon \in \mathbb{R}_{>0}$. Then, by definition of absolute continuity, there exists $\delta \in \mathbb{R}_{>0}$ such that, if $[c,d] \subseteq [a,b]$ is an interval for which $d - c < \delta$, then $|f(d) - f(c)| < \epsilon$. In particular, if $y \in \mathsf{B}(\delta, x) \cap I$, then $|f(y) - f(x)| < \epsilon$, giving continuity of f at x. Now let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta \in \mathbb{R}_{>0}$ have the property that for any family $((a_j,b_j))_{j\in\{1,\dots,k\}}$ of disjoint intervals for which

$$\sum_{j=1}^{k} |b_j - a_j| < \delta,$$

we have

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < \epsilon.$$

Now let P be a partition of $[a,b]$ for which $|P| < \delta$, and let $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$. Noting that $((x_{j-1}, x_j))_{j\in\{1,\dots,k\}}$ is a finite family of disjoint intervals, we have

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})| < k\epsilon.$$

Since this holds for any partition P for which $|P| < \delta$, and since the expression

$$\sum_{j=1}^{k} |f(x_j) - f(x_{j-1})|$$

is monotonically increasing as a function of $|P|$, it follows that

$$\mathrm{TV}(f) = \sup\Big\{ \sum_{j=1}^{l} |f(x_j) - f(x_{j-1})| \;\Big|\; (x_0, x_1, \dots, x_l) = \mathrm{EP}(P),\ P \in \mathrm{Part}([a,b]) \Big\} \le k\epsilon,$$

showing that f has bounded variation.

The result for general intervals follows directly from the result for compact intervals, along with the definition of local absolute continuity. ■

The converse of the preceding result is generally not true, as the following example illustrates.

5.9.25 Example (A continuous function of bounded variation that is not absolutely continuous) We consider the Cantor function $f_C\colon [0,1] \to \mathbb{R}$ of Example 3.2.27. We have shown that f_C is continuous, and since it is monotonically increasing, it is necessarily of bounded variation by Theorem 3.3.3. We claim, nonetheless, that f_C is not locally absolutely continuous. To see this, let $\delta \in \mathbb{R}_{>0}$. Recall from Example 2.5.39 that C is the intersection of a family $(C_k)_{k\in\mathbb{Z}_{>0}}$ of sets for which each of the sets C_k is a collection of 2^k disjoint closed intervals of length 3^{-k}. Therefore, since the total lengths of the intervals comprising C_k (i.e., $\lim_{k\to\infty} 2^k 3^{-k}$) goes to zero as k goes to infinity, there exists $N \in \mathbb{Z}_{>0}$ such that we can cover C_N with a finite family, say $((a_j, b_j))_{j\in\{1,\dots,2^N\}}$, of disjoint open intervals for which

$$\sum_{j=1}^{2^N} |b_j - a_j| < \delta.$$

Now note that since C is closed, $[0,1] \setminus C$ is open, and so, by Proposition 2.5.6, is a countable union of open intervals. By construction, f_C is constant on each of these open intervals. Since $C \subseteq C_N$, it follows that $[0,1] \setminus C_N \subseteq [0,1] \setminus C$ and so $[0,1]\setminus C_N$ is itself a countable (in fact, finite) collection of open intervals, each having the property that f_C is constant when restricted to it. Since f_C is monotonically increasing and continuous, it then follows that

$$\sum_{j=1}^{2^N} |f_C(b_j) - f_C(a_j)| = f(1) - f(0) = 1.$$

Since this conclusion is independent of $\delta \in \mathbb{R}_{>0}$, we therefore are forced to deduce that f_C is not absolutely continuous. •

This example illustrates that there is a "gap" between the notion of absolute continuity and the notion of continuous and bounded variation. It is perhaps not immediately clear why we should care about this. The reason we will care comes about in ***missing stuff*** where we shall see that any function of bounded variation is a sum of three functions, one being a saltus function, one being absolutely continuous, and the other being continuous, but not absolutely continuous (such functions we will call "singular").

Locally absolutely continuous functions, by virtue of also being of locally bounded variation, are almost everywhere differentiable. The next result we state provides us with a large collection of locally absolutely continuous functions based on their differentiability. The result also strengthens Proposition 3.3.14 where the hypotheses are the same, but here we draw the sharper conclusion of absolute continuity, not just bounded variation.

5.9.26 Proposition (Nice differentiable functions are locally absolutely continuous) *If* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval and if* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *is a differentiable function having the property that* f' *is locally bounded, then* f *is locally absolutely continuous.*

Proof Clearly it suffices to consider the case where $I = [a, b]$. Let $M \in \mathbb{R}_{>0}$ have the property that $|f'(x)| < M$ for each $x \in [a, b]$. Then, for $\epsilon \in \mathbb{R}_{>0}$ take $\delta = \frac{\epsilon}{M}$ and note that, if $((a_j, b_j))_{j\in\{1,\dots,k\}}$ is a finite family of disjoint open intervals satisfying

$$\sum_{j=1}^{k} |b_j - a_j| < \epsilon,$$

then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| = \sum_{j=1}^{k} |f'(c_j)(b_j - a_j)| < \epsilon,$$

where $c_j \in (a_j, b_j)$, $j \in \{1, \dots, k\}$, are as asserted by the Mean Value Theorem. ■

The boundedness of the derivative in the preceding result is essential, as Example 3.3.15 shows.

Let us next consider how absolutely continuous functions behave under the standard algebraic operations on functions. First we consider the standard algebraic operations.

5.9.27 Proposition (Addition and multiplication, and local absolute continuity) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval and let* $\mathrm{f}, \mathrm{g}\colon \mathrm{I} \to \mathbb{R}$ *be locally absolutely continuous. Then the following statements hold:*

(i) $\mathrm{f} + \mathrm{g}$ *is locally absolutely continuous;*

(ii) fg *is locally absolutely continuous;*

(iii) if additionally there exists $\alpha \in \mathbb{R}_{>0}$ *such that* $|\mathrm{g}(\mathrm{x})| \geq \alpha$ *for all* $\mathrm{x} \in \mathrm{I}$*, then* $\frac{\mathrm{f}}{\mathrm{g}}$ *is locally absolutely continuous.*

Proof Throughout the proof we suppose, without loss of generality, that $I = [a, b]$ is a compact interval.

(i) For $\epsilon \in \mathbb{R}_{>0}$ let $\delta \in \mathbb{R}_{>0}$ have the property that, if $((a_j, b_j))_{j\in\{1,\dots,k\}}$ is a finite family of disjoint open intervals satisfying

$$\sum_{j=1}^{k} |b_j - a_j| < \delta,$$

then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < \tfrac{\epsilon}{2}, \quad \sum_{j=1}^{k} |g(b_j) - g(a_j)| < \tfrac{\epsilon}{2}.$$

Then, again for any finite collection $((a_j, b_j))_{j\in\{1,\dots,k\}}$ of disjoint open intervals satisfying

$$\sum_{j=1}^{k}|b_j - a_j| < \delta,$$

we have

$$\sum_{j=1}^{k}|(f+g)(b_j) - (f+g)(a_j)| \le \sum_{j=1}^{k}|f(b_j) - f(a_j)| + \sum_{j=1}^{k}|g(b_j) - g(a_j)| < \epsilon,$$

using the triangle inequality.

(ii) Let

$$M_f = \sup\{|f(x)| \mid x \in [a,b]\}, \quad M_g = \sup\{|g(x)| \mid x \in [a,b]\}.$$

Let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta \in \mathbb{R}_{>0}$ have the property that, if $((a_j, b_j))_{j\in\{1,\dots,k\}}$ is a finite family of disjoint open intervals satisfying

$$\sum_{j=1}^{k}|b_j - a_j| < \delta,$$

then

$$\sum_{j=1}^{k}|f(b_j) - f(a_j)| < \frac{\epsilon}{2M_f}, \quad \sum_{j=1}^{k}|g(b_j) - g(a_j)| < \frac{\epsilon}{2M_g}.$$

Then, for any finite collection $((a_j, b_j))_{j\in\{1,\dots,k\}}$ of disjoint open intervals satisfying

$$\sum_{j=1}^{k}|b_j - a_j| < \delta,$$

we compute

$$\begin{aligned}\sum_{j=1}^{k}|f(b_j)g(b_j) - f(a_j)g(a_j)| &\le \sum_{j=1}^{k}|f(b_j)g(b_j) - f(a_j)g(b_j)| \\ &\quad + \sum_{j=1}^{k}|f(a_j)g(b_j) - f(a_j)g(a_j)| \\ &\le \sum_{j=1}^{k}M_g|f(b_j) - f(a_j)| + \sum_{j=1}^{k}M_f|g(b_j) - g(a_j)| \\ &< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,\end{aligned}$$

giving the result.

(iii) Let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta \in \mathbb{R}_{>0}$ have the property that, if $((a_j, b_j))_{j\in\{1,\dots,k\}}$ is a finite collection of disjoint open intervals satisfying

$$\sum_{j=1}^{k}|b_j - a_j| < \delta,$$

then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < \alpha^2 \epsilon.$$

Then, for any finite collection $((a_j, b_j))_{j \in \{1,\dots,k\}}$ of disjoint open intervals satisfying

$$\sum_{j=1}^{k} |b_j - a_j| < \delta,$$

we compute

$$\sum_{j=1}^{k} \Big| \frac{1}{g(b_j)} - \frac{1}{g(a_j)} \Big| = \sum_{j=1}^{k} \Big| \frac{g(a_j) - g(b_j)}{g(b_j) g(a_j)} \Big| \le \sum_{j=1}^{k} \Big| \frac{g(b_j) - g(a_j)}{\alpha^2} \Big| < \epsilon.$$

Thus $\frac{1}{g}$ is locally absolutely continuous, and this part of the result follows from part (ii). ■

Next let us show that local absolute continuity for a function on an interval can be determined by breaking the interval into parts, and determining local absolute continuity on each.

5.9.28 Proposition (Local absolute continuity on disjoint subintervals) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval and let* $\mathrm{I} = \mathrm{I}_1 \cup \mathrm{I}_2$*, where* $\mathrm{I}_1 \cap \mathrm{I}_2 = \{\mathrm{c}\}$*, where* c *is the right endpoint of* I_1 *and the left endpoint of* I_2*. Then* f *is locally absolutely continuous if and only if* $\mathrm{f}|\mathrm{I}_1$ *and* $\mathrm{f}|\mathrm{I}_2$ *are locally absolutely continuous.*

Proof It suffices to consider the case where $I = [a, c]$, $I_1 = [a, c]$, and $I_2 = [c, b]$.

First suppose that f is absolutely continuous and, for $\epsilon \in \mathbb{R}_{>0}$, choose $\delta \in \mathbb{R}_{>0}$ such that, if $((a_j, b_j))_{j \in \{1,\dots,k\}}$ is a finite family of disjoint open intervals satisfying

$$\sum_{j=1}^{k} |b_j - a_j| < 2\delta.$$

then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < \epsilon.$$

Then let $((a_j, c_j))_{j \in \{1,\dots,k_1\}}$ and $((d_j, b_j))_{j \in \{1,\dots,k_2\}}$ be finite families of disjoint open subintervals of $[a, c]$ and $[a, c]$, respectively, satisfying

$$\sum_{j=1}^{k_1} |c_j - a_j| < \delta, \quad \sum_{j=1}^{k_2} |b_j - d_j| < \delta.$$

Then $((a_j, c_j))_{j \in \{1,\dots,k_1\}} \cup ((d_j, b_j))_{j \in \{1,\dots,k_2\}}$ is a finite collection of disjoint open subintervals of $[a, b]$ satisfying

$$\sum_{j=1}^{k_1} |c_j - a_j| + \sum_{j=1}^{k_2} |b_j - d_j| < 2\delta.$$

Therefore,

$$\sum_{j=1}^{k_1} |f(c_j) - f(a_j)| + \sum_{j=1}^{k_2} |f(b_j) - f(d_j)| < \epsilon,$$

implying that

$$\sum_{j=1}^{k_1} |f(c_j) - f(a_j)| < \epsilon, \quad \sum_{j=1}^{k_2} |f(b_j) - f(d_j)| < \epsilon.$$

Thus $f|[a,c]$ and $f|[c,b]$ are absolutely continuous.

Now suppose that $f|[a,c]$ and $f|[c,b]$ are absolutely continuous. Let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta \in \mathbb{R}_{>0}$ be chosen such that, if $((a_j, c_j))_{j\in\{1,\dots,k_1\}}$ and $((d_j, b_j))_{j\in\{1,\dots,k_2\}}$ are finite collections of disjoint open subintervals of $[a,c]$ and $[c,b]$, respectively, satisfying

$$\sum_{j=1}^{k_1} |c_j - a_j| < \delta, \quad \sum_{j=1}^{k_2} |b_j - d_j| < \delta,$$

then

$$\sum_{j=1}^{k_1} |f(c_j) - f(a_j)| < \tfrac{\epsilon}{2}, \quad \sum_{j=1}^{k_2} |f(b_j) - f(d_j)| < \tfrac{\epsilon}{2}.$$

Now let $((a_j, b_j))_{j\in\{1,\dots,k\}}$ be a finite collection of disjoint subintervals of $[a,b]$ satisfying

$$\sum_{j=1}^{k} |b_j - a_j| < \delta.$$

If $c \in (a_{j_0}, b_{j_0})$ for some $j_0 \in \{1,\dots,k\}$, then define the collection of disjoint open intervals

$$(((a_j, b_j))_{j\in\{1,\dots,k\}} \setminus ((a_{j_0}, b_{j_0}))) \cup ((a_{j_0}, c), (c, b_{j_0})),$$

i.e., split the interval containing c into two intervals. Denote this collection of disjoint open intervals by $((\tilde{a}_j, \tilde{b}_j))_{j\in\{1,\dots,\tilde{k}\}}$. If c is not contained in any of the intervals $((a_j, b_j))_{j\in\{1,\dots,k\}}$, then denote $((\tilde{a}_j, \tilde{b}_j))_{j\in\{1,\dots,\tilde{k}\}} = ((a_j, b_j))_{j\in\{1,\dots,k\}}$. Note that

$$\sum_{j=1}^{\tilde{k}} |\tilde{b}_j - \tilde{a}_j| < \delta.$$

This new collection of disjoint open intervals is then the union of two collections of disjoint open intervals, $((\tilde{a}_j, \tilde{c}_j))_{j\in\{1,\dots,k_1\}}$ and $((\tilde{d}_j, \tilde{b}_j))_{j\in\{1,\dots,k_2\}}$, the first being subintervals of $[a,c]$ and the second being subintervals of $[c,b]$. These collections satisfy

$$\sum_{j=1}^{k_1} |\tilde{c}_j - \tilde{a}_j| < \delta, \quad \sum_{j=1}^{k_2} |\tilde{b}_j - \tilde{d}_j| < \delta,$$

and so we have

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| \le \sum_{j=1}^{\tilde{k}} |f(\tilde{b}_j) - f(\tilde{a}_j)| = \sum_{j=1}^{k_1} |f(\tilde{c}_j) - f(\tilde{a}_j)| + \sum_{j=1}^{k_2} |f(\tilde{b}_j) - f(\tilde{d}_j)| < \epsilon,$$

which shows that f is absolutely continuous. ■

Next we show that one of the standard operations on functions does *not* respect absolute continuity.

5.9.29 Example (Compositions of locally absolutely continuous functions need not be locally absolutely continuous) In Example 3.3.16 we gave two functions of bounded variation whose composition was not a function of bounded variation. In fact, the functions we used were not only of bounded variation, but absolutely continuous. These functions, therefore, show that the composition of absolutely continuous functions may not be an absolutely continuous function.

Let us show that the functions in question are, in fact, absolutely continuous. Recall that the functions $f, g\colon [-1,1] \to \mathbb{R}$ were given by $f(x) = x^{1/3}$ and by

$$g(x) = \begin{cases} x^3(\sin\frac{1}{x})^3, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

That g is absolutely continuous follows from Proposition 5.9.26 since we showed in Example 3.3.16 that g was of class C^1. It then only remains to show that f is absolutely continuous. Let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \frac{\epsilon^3}{4}$. Now let $((a_j, b_j))_{j\in\{1,\dots,k\}}$ be a finite collection of open intervals satisfying

$$\sum_{j=1}^{k} |b_j - a_j| < \delta.$$

Let $\ell \leq 2$ and let $[a,b] \subseteq [-1,1]$ be an interval of length ℓ. One can easily see that, if one fixes the length of the interval at ℓ, then the quantity $|f(b) - f(a)|$ is maximum when one takes $a = -\frac{\ell}{2}$ and $b = \frac{\ell}{2}$. From this it follows that

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < |f(\tfrac{\delta}{2}) - f(-\tfrac{\delta}{2})| = \epsilon.$$

Thus f is absolutely continuous, as desired. •

5.9.7 The Fundamental Theorem of Calculus for the Lebesgue integral

In this section we explore the Fundamental Theorem of Calculus that is associated with the Lebesgue integral. As we shall see, it is here that the notion of absolute continuity comes up in a natural way.

Before we state the main result,

5.9.30 Lemma (Locally absolutely continuous functions with a.e. zero derivative) *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be an interval and let* $\mathrm{f}\colon \mathrm{I} \to \mathbb{R}$ *be locally absolutely continuous and having the property that the set*

$$\{\mathrm{x} \in \mathrm{I} \mid \mathrm{f} \textit{ is not differentiable at } \mathrm{x}\} \cap \{\mathrm{x} \in \mathrm{I} \mid \mathrm{f}'(\mathrm{x}) \neq 0\}$$

has measure zero. Then there exists $\mathrm{c} \in \mathbb{R}$ *such that* $\mathrm{f}(\mathrm{x}) = \mathrm{c}$ *for all* $\mathrm{x} \in \mathrm{I}$.

Proof Consider an interval $[a,b] \subseteq I$. Let

$$E = \{x \in I \mid f \text{ is not differentiable at } x\} \cap \{x \in I \mid f'(x) \neq 0\}.$$

For $\epsilon \in \mathbb{R}_{>0}$ choose $\eta \in \mathbb{R}_{>0}$ such that, if $((a_j, b_j))_{j\in\{1,\dots,k\}}$ is a finite collection of disjoint intervals having the property that

$$\sum_{j=1}^{k} |b_j - a_j| < \eta,$$

then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| < \epsilon.$$

Let $((c_\alpha, d_\alpha))_{\alpha\in A}$ be a countable collection of open intervals satisfying

$$E \subseteq \bigcup_{\alpha\in A}(c_\alpha, d_\alpha)$$

and

$$\sum_{\alpha\in A} |d_\alpha - c_\alpha| < \eta.$$

Now define $\delta\colon [a,b] \to \mathbb{R}_{>0}$ according to the following:

1. if $x \in E$ take $\delta(x)$ such that $\mathsf{B}(\delta(x), x) \cap [a,b] \subseteq (c_\alpha, d_\alpha)$ for some $\alpha \in A$;
2. if $x \notin E$ take $\delta(x)$ such that $|f(y) - f(x)| < \epsilon|y - x|$ for $y \in \mathsf{B}(\delta(x), x) \cap [a,b]$.

Now let $((c_1, I_1), \dots, (c_k, I_k))$ be a δ-fine tagged partition and write $\{1,\dots,k\} = K_1 \overset{\circ}{\cup} K_2$ where

$$K_1 = \{j \in \{1,\dots,k\} \mid c_j \in E\}, \quad K_2 = \{j \in \{1,\dots,k\} \mid c_j \notin E\}.$$

We then compute, denoting $\mathrm{EP}(P) = (x_0, x_1, \dots, x_k)$,

$$\begin{aligned} |f(b) - f(a)| &= \Big|\sum_{j=1}^{k} (f(x_j) - f(x_{j-1}))\Big| \\ &\le \sum_{j\in K_1} |f(x_j) - f(x_{j-1})| + \sum_{j\in K_2} |f(x_j) - f(x_{j-1})| \\ &\le \epsilon + \sum_{j\in K_2} \epsilon(x_j - x_{j-1}) \le \epsilon(1 + b - a). \end{aligned}$$

This shows that $|f(b) - f(a)|$ can be made arbitrarily small, and so gives the result since a and b are arbitrary. ■

The main result in this section is the following.

5.9.31 Theorem (The Fundamental Theorem of Calculus for the Lebesgue integral) *For an interval* $\mathrm{I} \subseteq \mathbb{R}$ *the following statements hold:*

(i) a function $\mathrm{F}\colon \mathrm{I} \to \mathbb{R}$ *defined on an interval* I *is locally absolutely continuous if and only if there exists* $\mathrm{f} \in \mathsf{L}^{(1)}_{\mathrm{loc}}(\mathrm{I}; \mathbb{R})$ *and* $\mathrm{x}_0 \in \mathrm{I}$ *such that*

$$\mathrm{F}(\mathrm{x}) = \mathrm{F}(\mathrm{x}_0) + \int_{\mathrm{x}_0}^{\mathrm{x}} \mathrm{f}(\xi)\,\mathrm{d}\xi,$$

where we adopt the convention that if $x < x_0$ *we have*

$$\int_{x_0}^{x} g(\xi)\,d\xi = -\int_{x}^{x_0} g(\xi)\,d\xi;$$

(ii) if $x_0 \in I$*, if* $f \in \mathsf{L}^{(1)}_{\mathrm{loc}}(I;\mathbb{R})$*, and define* $F\colon I \to \mathbb{R}$ *by*

$$F(x) = \int_{x_0}^{x} f(\xi)\,d\xi,$$

then F *is differentiable for almost every* $x \in I$ *and* $F'(x) = f(x)$ *for almost every* $x \in I$*.*

Proof (i) We first consider the case when I is compact: $I = [a,b]$.

First suppose that

$$F(x) = F(x_0) + \int_{x_0}^{x} f(\xi)\,d\xi$$

for $x_0 \in [a,b]$ and $f \in \mathsf{L}^{(1)}([a,b];\mathbb{R})$. Note that, by Proposition 5.7.22, we have

$$\begin{aligned} F(x) &= F(x_0) + \int_a^x f(\xi)\,d\xi - \int_a^{x_0} f(\xi)\,d\xi \\ &= F(x_0) + \int_a^x f(\xi)\,d\xi + (F(a) - F(x_0) = F(a) + \int_a^x f(\xi)\,d\xi. \end{aligned}$$

Thus we can take $x_0 = a$ without loss of generality. First assume that f is nonnegative-valued. For $k \in \mathbb{Z}_{>0}$ define

$$f_k(x) = \begin{cases} f(x), & f(x) \le k \\ k, & \text{otherwise.} \end{cases}$$

Note that f_k is bounded and that for each $x \in [a,b]$ we have $\lim_{k\to\infty} f_k(t) = f(t)$. Therefore, the Monotone Convergence Theorem asserts that

$$\lim_{k\to\infty} \int_a^b (f(x) - f_k(x))\,dx = 0.$$

Now let $\epsilon \in \mathbb{R}_{>0}$. Choose $N \in \mathbb{Z}_{>0}$ such that

$$\int_a^b (f(x) - f_k(x))\,dx < \frac{\epsilon}{2}, \qquad k \ge N.$$

Letting $\delta = \frac{\epsilon}{2N}$ and letting $((a_j,b_j))_{j\in\{1,\dots,n\}}$ be any finite family of nonoverlapping intervals in $[a,b]$ satisfying

$$\sum_{j=1}^{n} |b_j - a_j| < \delta,$$

we have

$$\int_A f(x)\,dx = \int_A (f(x) - f_N(x))\,dx + \int_A f_N(x)\,dx \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

where A denotes the union of the intervals $((a_j, b_j))_{j\in\{1,\dots,n\}}$. Note that since f is nonnegative, it follows that F is monotonically increasing. Thus

$$\sum_{j=1}^{n} |F(b_j) - F(a_j)| = \sum_{j=1}^{n} (F(b_j) - F(a_j)).$$

Given the definition of F we thus have

$$\sum_{j=1}^{n} |F(b_j) - F(a_j)| = \int_A f(x)\,\mathrm{d}x < \epsilon,$$

and we conclude that F is absolutely continuous.

If f is not nonnegative-valued, then we write $f = f_+ - f_-$ where f_+ and f_- are nonnegative. Our arguments above show that the functions

$$x \mapsto \int_a^x f_+(\xi)\,\mathrm{d}\xi, \quad x \mapsto \int_a^x f_-(\xi)\,\mathrm{d}\xi$$

are absolutely continuous. Therefore, since

$$F(x) = F(a) + \int_a^x f_+(\xi)\,\mathrm{d}\xi - \int_a^x f_-(\xi)\,\mathrm{d}\xi,$$

it follows that F is the sum of three absolutely continuous functions (a constant function is trivially absolutely continuous) and so F is itself absolutely continuous by Proposition 5.9.27.

Now suppose that F is absolutely continuous, and so of bounded variation by Proposition 5.9.24. Now, by part (ii) of Theorem 3.3.3, write $F = F_+ - F_-$ for monotonic functions F_+ and F_-. By part (vi) of Theorem 3.3.3 the derivative of F exists almost everywhere and we then have

$$F'(x) = F'_+(x) - F'_-(x) \qquad \Longrightarrow \qquad |F'(x)| \le |F'_+(x)| + |F'_-(x)|$$

for almost every $x \in [a, b]$. Therefore we have

$$\int_a^b |F'(x)|\,\mathrm{d}x \le F_+(b) + F_-(b) - F_+(a) - F_-(a),$$

implying that $F' \in \mathsf{L}^{(1)}([a, b]; \mathbb{R})$. Note that the function

$$x \mapsto \int_a^x F'(\xi)\,\mathrm{d}\xi$$

is now absolutely continuous by our arguments from the first part of the proof, so that the function

$$x \mapsto F(x) - \int_a^x F'(\xi)\,\mathrm{d}\xi$$

is also absolutely continuous by Proposition 5.9.27. This function also has derivative zero, and the result now follows by Lemma 5.9.30.

Now suppose that I is an arbitrary interval. We first suppose that

$$F(x) = F(x_0) + \int_{x_0}^{x} f(\xi)\,\mathrm{d}\xi$$

for some $x_0 \in I$ and some locally integrable function f. Let $[a,b] \subseteq I$ be a compact subinterval. As we determined in the first part of the proof, we have

$$F(x) = F(a) + \int_{a}^{x} f(\xi)\,\mathrm{d}\xi,$$

from which we conclude that $F|[a,b]$ is absolutely continuous, since we proved have already proved the theorem for compact intervals. It then follows that F is locally absolutely continuous since this can be done for any compact subinterval. Conversely, suppose that F is locally absolutely continuous and let $x_0 \in I$. Let $x \in I$, supposing that $x > x_0$. Note that, since $F|[x_0,x]$ is absolutely continuous, the first part of the proof allows us to conclude that

$$F(x) = F(x_0) + \int_{x_0}^{x} f(\xi)\,\mathrm{d}\xi.$$

If $x < x_0$ we have that $F|[x,x_0]$ is absolutely continuous and so we can write

$$F(x_0) = F(x) + \int_{x}^{x_0} f(\xi)\,\mathrm{d}\xi,$$

and the theorem follows by a rearrangement of this equation, using the stated convention for integrals whose lower limit exceeds the upper limit.

(ii) We first prove a technical lemma from which this part of the theorem will follow.

1 Lemma *If* $\mathsf{A} \subseteq \mathbb{R}$ *then*

$$\lim_{\beta\downarrow 0} \frac{\lambda(\mathsf{A} \cap (\mathsf{x}, \mathsf{x}+\beta))}{\beta} = \lim_{\alpha\downarrow 0} \frac{\lambda(\mathsf{A} \cap (\mathsf{x}-\alpha, \mathsf{x}))}{\alpha} = \lim_{\alpha,\beta\downarrow 0} \frac{\lambda(\mathsf{A} \cap (\mathsf{x}-\alpha, \mathsf{x}+\beta))}{\alpha+\beta} = 1$$

for almost every $\mathsf{x} \in \mathsf{A}$. *If we additionally have* $\mathsf{A} \in \mathscr{L}(\mathbb{R})$ *then the above limits are equal to zero for almost every* $\mathsf{x} \in \mathbb{R} \setminus \mathsf{A}$.

Proof First suppose that A is bounded so that $\lambda^*(A) < \infty$. By definition of Lebesgue outer measure, for $k \in \mathbb{Z}_{>0}$ there exists a countable collection $((a_{k,j}, b_{k,j}))_{j\in\mathbb{Z}_{>0}}$ of open intervals such that

$$\sum_{j=1}^{\infty} |b_{k,j} - a_{k,j}| - 2^{-k} < \lambda^*(A).$$

If we define $U''_k = \cup_{j=1}^{\infty}(a_{k,j}, b_{k,j})$ then we have

$$\lambda(U''_k) - 2^{-k} \le \sum_{j=1}^{\infty} |b_{k,j} - a_{k,j}| - 2^{-k} < \lambda^*(A).$$

Then define $U'_m = \cap_{k=1}^{m} U''_k$ so that $U'_{m+1} \subseteq U'_m$, $m \in \mathbb{Z}_{>0}$, and

$$\lambda(U'_m) - 2^{-m} \le \lambda(U''_m) - 2^{-m} < \lambda^*(A).$$

Finally, let (a, b) be such that $A \subseteq (a, b)$ and define $U_k = U'_k \cap (a, b)$, $k \in \mathbb{Z}_{>0}$. Then $A \subseteq \cap_{k \in \mathbb{Z}_{>0}} U_k$, $U_{k+1} \subseteq U_k$, $k \in \mathbb{Z}_{>0}$, and $\lambda(U_k) - 2^{-k} < \lambda^*(A)$, $k \in \mathbb{Z}_{>0}$.

Now define $f_k \colon \mathbb{R} \to \mathbb{R}$, $k \in \mathbb{Z}_{>0}$, and $f \colon \mathbb{R} \to \mathbb{R}$ by

$$f_k(x) = \lambda(U_k \cap (a, x)), \quad f(x) = \lambda(A \cap (a, x)).$$

Since U_k is open for $k \in \mathbb{Z}_{>0}$, if $x \in U_k$ and if $\epsilon \in \mathbb{R}_{>0}$ is such that $(x - \epsilon, x + \epsilon) \subseteq U_k$ we have

$$\frac{f_k(x+\epsilon) - f_k(x)}{\epsilon} = \frac{f_k(x) - f_k(x-\epsilon)}{\epsilon} = 1.$$

Thus $f_k|U_k$ is differentiable with derivative 1.

Let $x_1, x_2 \in \mathbb{R}$ with $a \leq x_1 < x_2$. Then, for each $k \in \mathbb{Z}_{>0}$,

$$\begin{aligned} f_k(x_2) - f(x_2) - (f_k(x_1) - f(x_1)) &= \lambda(U_k \cap [x_1, x_2)) - \lambda(A \cap (a, x_2)) + \lambda(A \cap (a, x_1)) \\ &= \lambda(U_k \cap [x_1, x_2)) - \lambda(A \cap [x_1, x_2)) \geq 0 \end{aligned}$$

by monotonicity of Lebesgue measure. This shows that the function $f_k - f$ is monotonically increasing for each $k \in \mathbb{Z}_{>0}$. We also have

$$f_k(b) - f(b) = \lambda(U_k) - \lambda(A) < 2^{-k}$$

which gives

$$\sum_{k=1}^{\infty} (f_k(x) - f(x)) \leq \sum_{k=1}^{\infty} (f_k(b) - f(b)) \leq \sum_{k=1}^{\infty} 2^{-k} < \infty,$$

using Example 2.4.2–**??**. If we define

$$g(x) = \sum_{k=1}^{\infty} (f_k(x) - f(x)),$$

then, by Theorems 3.2.26 and 3.5.25, g is almost everywhere differentiable and

$$g'(x) = \sum_{k=1}^{\infty} (f'_k(x) - f'(x))$$

for almost every $x \in (a, b)$. Since g is monotonically increasing, $g'(x)$ is finite for almost every $x \in [a, b]$. This gives

$$\sum_{k=1}^{\infty} (f'_k(x) - f'(x)) < \infty$$

for almost every $x \in (a, b)$. Since $f'_k(x) - f'(x) \geq 0$ for almost every $x \in (a, b)$ we must have

$$\lim_{k \to \infty} f'_k(x) - f'(x) = 0$$

for almost every $x \in (a, b)$. Let $N \subseteq (a, b)$ be the set of points on which the above limit does not hold, so $\lambda^*(N) = 0$. Let $x \in (\cap_{k \in \mathbb{Z}_{>0}} U_k) - N$. Then $f'_k(x) = 1$ for every $k \in \mathbb{Z}_{>0}$ and so $f'(x) = 1$. Thus $f'(x) = 1$ for $x \in A - N$, giving the first assertion of the lemma in the case when A is bounded. If A is not bounded then we can write A as a countable union of bounded sets: $A = \cup_{j \in \mathbb{Z}_{>0}} A_j$. Let $N_j \subseteq A_j$ be the subset of A_j where the limits in the first assertion of the theorem do not have the value 1. Then the limits in the first assertion of the theorem hold for all $x \in A \setminus \cup_{j\in\mathbb{Z}_{>0}} N_j$. Since $\cup_{j\in\mathbb{Z}_{>0}} N_j$ has measure zero by Exercise 2.5.9, the first part of the theorem is proved.

For the second assertion, if A is measurable then we have

$$\alpha + \beta = \lambda((x-\alpha, x+\beta)) = \lambda(A \cap (x-\alpha, x+\beta)) + \lambda((\mathbb{R} \setminus A)(x-\alpha, x+\beta)).$$

Thus

$$1 = \frac{\lambda(A \cap (x-\alpha, x+\beta))}{\alpha+\beta} + \frac{\lambda((\mathbb{R} \setminus A)(x-\alpha, x+\beta))}{\alpha+\beta},$$

and taking the limit as α and β decrease to zero gives

$$\lim_{\alpha,\beta\downarrow 0} \frac{\lambda((\mathbb{R} \setminus A)(x-\alpha, x+\beta))}{\alpha+\beta},$$

using the fact that the first part of the proof has been proved. ▼

Proceeding with the proof of the theorem, first consider the case when $I = [a, b]$ and

$$F(x) = \int_a^x f(\xi)\,\mathrm{d}\xi;$$

the lower limit can be taken to be a as we saw in the first part of the proof. We first consider the case where f is a finite nonnegative simple function,

$$f(x) = \sum_{j=1}^k a_j \chi_{A_j}(x).$$

Then, by linearity of the integral,

$$F(x) = \int_a^x f(\xi)\,\mathrm{d}\xi = \sum_{j=1}^k a_j \lambda(A_j \cap (a, x)).$$

By the lemma it follows that F is differentiable for almost every $x \in (a, b)$ and that $F'(x) = f(x)$ for almost every $x \in (a, b)$.

Now suppose that f is a nonnegative simple function and let $(g_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of nonnegative simple functions as in part (ii) of Proposition 5.6.39. For $j \in \mathbb{Z}_{>0}$ define

$$G_j(x) = \int_a^x g_j(\xi)\,\mathrm{d}\xi.$$

By the Monotone Convergence Theorem,

$$\begin{aligned} F(x) = \int_a^x f(\xi)\,\mathrm{d}\xi = \lim_{j\to\infty} \int_a^x g_j(\xi)\,\mathrm{d}\xi = \lim_{j\to\infty} G_j(x) \\ = G_1(x) + \sum_{j=1}^{\infty} (G_{j+1}(x) - G_j(x)) \end{aligned}$$

for every $x \in [a, b]$. Note that for each $j \in \mathbb{Z}_{>0}$ the functions G_j and $G_{j+1} - G_j$ are monotonically increasing, being the indefinite integrals of nonnegative functions. Therefore, we can apply Theorem 3.5.25 to arrive at the equality

$$F'(x) = G_1'(x) + \sum_{j=1}^{\infty}(G_{j+1}'(x) - G_j'(x)) = \lim_{j\to\infty} G_j'(x)$$

for almost every $x \in (a, b)$. Since the theorem has been proved for nonnegative simple functions, we have

$$\lim_{j\to\infty} G_j'(x) = \lim_{j\to\infty} g_j(x) = f(x)$$

for almost every $x \in (a, b)$. Therefore, $F'(x) = f(x)$ for almost every $x \in (a, b)$.

Now let I be an arbitrary interval with $x_0 \in I$ and

$$F(x) = \int_{x_0}^{x} f(\xi)\,\mathrm{d}\xi.$$

Let $(I_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of bounded intervals all containing x_0 such that $I_j \subseteq I_{j+1}$, $j \in \mathbb{Z}_{>0}$, and such that $\cup_{j\in\mathbb{Z}_{>0}} I_j = I$ (make sure you understand why this is possible). By the arguments above, $F'(x) = f(x)$ for almost every $x \in I_j$ and for every $j \in \mathbb{Z}_{>0}$. Thus, if $N_j \subseteq I_j$ is the set of measure zero for which F' does not exist or, if it exists is not equal to $f(x)$, then $F'(x) = f(x)$ for all $x \in I \setminus \cup_{j\in\mathbb{Z}_{>0}} N_j$. Since $\lambda(\cup_{j\in\mathbb{Z}_{>0}} N_j) = 0$ by Exercise 2.5.9, the theorem follows. ■

In Example 3.4.31 we considered a collection of examples illustrating the Fundamental Theorem of Calculus for the Riemann integral. The examples where this version of the Fundamental Theorem applies still apply for the Lebesgue integral by virtue of Theorem 5.9.11. However, in Example 3.4.31 we saw an instance of a differentiable function on $[0, 1]$ that is everywhere differentiable and with bounded derivative, but the derivative is not Riemann integrable. This example is more satisfactory with the Lebesgue integral.

5.9.32 Example (The Fundamental Theorem of Calculus for the Lebesgue integral) The reader should go back and carefully read the construction of Example 3.4.31. The reader will see that the example is of a function $F\colon [0, 1] \to \mathbb{R}$ with the property that F is everywhere differentiable with a bounded derivative. However, F' is not Riemann integrable. By Proposition 5.9.26, however, F is absolutely continuous, and so F' is Lebesgue integrable by Theorem 5.9.31. •

One of the conclusions of the Fundamental Theorem of Calculus for the Lebesgue integral is that an absolutely continuous function is almost everywhere differentiable. As we saw in Proposition 5.9.24, absolutely continuous functions are continuous. One might speculate, then, that a characterisation of absolute continuity using continuity and the derivative might be possible. For example, here are some guesses, along with counterexamples.

1. *An absolutely continuous function is one that is continuous and differentiable almost everywhere*. This is false as seen by Example 3.3.15.

2. *An absolutely continuous function is one that is continuous, differentiable almost everywhere, and with integrable derivative.* This is false by virtue of Example 5.9.25.
3. *An absolutely continuous function is one that is differentiable almost everywhere.* This is false by virtue of Example 3.3.15.

However, there is the following result, which is sometimes enough to understand absolute continuity.

5.9.33 Theorem (A class of absolutely continuous functions) *If* $\mathrm{F}\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}$ *is*

(i) continuous,

(ii) differentiable at all but at most countable many points in $[\mathrm{a},\mathrm{b}]$*, and*

(iii) the function

$$\mathrm{f}(\mathrm{x}) = \begin{cases} \mathrm{F}'(\mathrm{x}), & \textit{the derivative exists,} \\ 0, & \textit{otherwise,} \end{cases}$$

is in $\mathsf{L}^{(1)}([\mathrm{a},\mathrm{b}];\mathbb{R})$*,*

then F *is absolutely continuous.*

Proof Our proof relies on the definition in Section **??** of lower semicontinuous functions. Note that, by Proposition **??**, lower semicontinuous functions are Borel measurable. With this notion recalled, we have the following lemma.

1 Lemma *If* $\mathrm{f} \in \mathsf{L}^{(1)}([\mathrm{a},\mathrm{b}];\overline{\mathbb{R}})$ *then, for each* $\epsilon \in \mathbb{R}_{>0}$*, there exists a lower semicontinuous* $\mathrm{g} \in \mathsf{L}^{(1)}([\mathrm{a},\mathrm{b}];(-\infty,\infty])$ *such that* $\mathrm{f}(\mathrm{x}) \leq \mathrm{g}(\mathrm{x})$ *for every* $\mathrm{x} \in [\mathrm{a},\mathrm{b}]$ *and*

$$\int_\mathrm{a}^\mathrm{b} \mathrm{g}(\mathrm{x})\,\mathrm{d}\mathrm{x} < \int_\mathrm{a}^\mathrm{b} \mathrm{f}(\mathrm{x})\,\mathrm{d}\mathrm{x} + \epsilon.$$

Proof Let $\epsilon \in \mathbb{R}_{>0}$. We first consider the case when f is nonnegative-valued. We let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of simple functions as in part (ii) of Proposition 5.6.39. Then

$$f(x) = \lim_{j\to\infty} f_j(x) = f_1(x) + \sum_{j=1}^{\infty}(f_{j+1}(x) - f_j(x)). \tag{5.21}$$

Let $j \in \mathbb{Z}_{>0}$ and write

$$f_j = \sum_{k=1}^{m} a_k\chi_{A_k}, \quad f_{j+1} = \sum_{l=1}^{n} b_l\chi_{B_l}.$$

For each $l \in \{1,\dots,n\}$ write

$$B_l = \cup_{k=1}^{n}(A_k \cap B_l).$$

If $A_k \cap B_l \neq \emptyset$ then, on $A_k \cap B_l$ the value of $f_{j+1} - f_j$ is $b_l - a_k \in \mathbb{R}_{>0}$. Thus $(f_{j+1} - f_j)|B_l$ is a nonnegative simple function. Since this is true for every l it follows that $f_{j+1} - f_j$ is a nonnegative simple function. Thus, by (5.21), f is an infinite sum of nonnegative simple functions. Thus we write

$$f = \sum_{k=1}^{\infty} a_k\chi_{A_k}$$

where the numbers $a_k \in \mathbb{R}_{>0}$ and the sets A_k, $k \in \mathbb{Z}_{>0}$, are not related to those above. For $k \in \mathbb{Z}_{>0}$ let U_k be an open set such that $A_k \subseteq U_k$ and such that

$$\lambda(U_k) < \lambda(A_k) + \tfrac{\epsilon}{a_k 2^k}.$$

Then

$$\sum_{k=1}^{\infty} a_k \lambda(U_k) < \sum_{k=1}^{\infty} a_k \lambda(A_k) + \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \sum_{k=1}^{\infty} a_k \lambda(A_k) + \epsilon,$$

where we use Example 2.4.2–**??**. By Example **??**–**??** each of the functions $a_k \chi_{U_k}$ is lower semicontinuous. Define

$$h_m(x) = \sum_{k=1}^{m} a_k \lambda(U_k)$$

and

$$h(x) = \sum_{k=1}^{\infty} a_k \lambda(U_k).$$

Then h_m is lower semicontinuous by ***missing stuff***, and, since

$$h(x) = \sup\{h_m(x) \mid m \in \mathbb{Z}_{>0}\},$$

h is also lower semicontinuous by Proposition **??**. We then have

$$\int_a^b h(x)\,\mathrm{d}x < \int_a^b f(x)\,\mathrm{d}x + \epsilon$$

and $f(x) \le h(x)$ for all $x \in [a, b]$.

Now suppose that $f \in \mathsf{L}^{(1)}([a, b]; \mathbb{R})$ and, for $k \in \mathbb{Z}_{>0}$, define

$$f_k(x) = \begin{cases} f(x), & f(x) > -k, \\ -k, & f(x) \le -k. \end{cases}$$

By the Dominated Convergence Theorem,

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{k\to\infty} \int_a^b f_k(x)\,\mathrm{d}x.$$

Now let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\int_a^b f_N(x)\,\mathrm{d}x < \int_a^b f(x)\,\mathrm{d}x - \frac{\epsilon}{2}.$$

Since $f_N + N\chi_{[a,b]}$ is nonnegative, from the first part of the proof there exists a lower semicontinuous function h such that $f_N(x) + N \le h(x)$ for every $x \in [a, b]$ and such that

$$\int_a^b h(x)\,\mathrm{d}x < \int_a^b (f_N(x) + N\chi_{[a,b]})\,\mathrm{d}x + \frac{\epsilon}{2}.$$

Define $g = h - N\chi_{[a,b]}$. Then $f(x) \le f_N(x) \le g(x)$ for every $x \in [a, b]$ and

$$\int_a^b g(x)\,\mathrm{d}x = \int_a^b (h(x) - N\chi_{[a,b]})\,\mathrm{d}x < \int_a^b f_N(x)\,\mathrm{d}x < \int_a^b f(x)\,\mathrm{d}x - \frac{\epsilon}{2},$$

as desired. ▼

2 Lemma *Let* $\mathrm{h}\colon [\mathrm{a},\mathrm{b}] \to \mathbb{R}$ *be continuous and let* $\mathrm{C} \subseteq [\mathrm{a},\mathrm{b}]$ *be countable. If, for each* $\mathrm{x} \in [\mathrm{a},\mathrm{b}) - \mathrm{C}$*, there exists* $\mathrm{r_x} \in \mathbb{R}_{>0}$ *such that* $\mathrm{h}(\mathrm{z}) > \mathrm{h}(\mathrm{x})$ *for each* $\mathrm{z} \in (\mathrm{x}, \mathrm{x} + \mathrm{r_x})$*, then* h *is monotonically increasing.*

Proof Suppose that h is continuous and that $x_1, x_2 \in [a,b]$ satisfy $x_1 < x_2$ and $h(x_1) > h(x_2)$. For $y \in (h(x_2), h(x_1))$ define

$$x_y = \sup\{x \in [x_1, x_2] \mid h(x) > y\}.$$

Then there exists a sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ in $[x_1, x_2]$ such that $x_j \leq x_y$, $j \in \mathbb{Z}_{>0}$, and such that $\lim_{j\to\infty} x_j = x_y$. By continuity of h, $\lim_{j\to\infty} h(x_j) = y$, using Theorem 3.1.3. We claim that, for any $r_y \in \mathbb{R}_{>0}$, there exists $z \in (x_y, x_y + r_y)$ such that $h(z) \leq h(x_y)$. Indeed, were this not so, then there would exist $z > x_y$ such that $h(z) > h(x_y) = y$, contradicting the definition of x_y. Since this construction can be made for every $y \in (h(x_2), h(x_1))$, this shows, therefore, that the complement to the set

$$\{x \in [a,b) \mid \text{ there exists } r_x \in \mathbb{R}_{>0} \text{ such that } h(z) > h(x) \text{ for each } z \in (x, x + r_x)\}$$

is not countable, which give the lemma. ▼

Proceeding with the proof, let $\epsilon \in \mathbb{R}_{>0}$. Denote by $C \subseteq [a,b]$ the countable subset at whose points F is not differentiable. By Lemma 1 let $h\colon [a,b] \to (-\infty, \infty]$ be lower semicontinuous and such that $f(t) \leq h(t)$ for $t \in [a,b] \setminus C$ and such that

$$\int_a^b h(x)\,\mathrm{d}x < \int_a^b f(x)\,\mathrm{d}x + \frac{\epsilon}{2}.$$

Then, if we define $g = h + \frac{\epsilon}{2(b-a)}$, then $f(t) < g(t)$ for $t \in [a,b] \setminus C$ and

$$\int_a^b g(x)\,\mathrm{d}x < \int_a^b f(x)\,\mathrm{d}x + \epsilon.$$

Let $G\colon [a,b] \to \mathbb{R}$ be defined by

$$G(x) = F(a) + \int_a^x g(\xi)\,\mathrm{d}\xi$$

Let $x \in [a,b)$. Since g is lower semicontinuous, for each $\eta \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that, if $x' \in [x, x+\delta]$, we have $g(x') > g(x) - \eta$. Then, for any $y \in [x, x+\delta]$ we have

$$\begin{aligned} G(y) - G(x) = F(a) + \int_a^y g(\xi)\,\mathrm{d}\xi - F(a) - \int_a^x g(\xi)\,\mathrm{d}\xi &= \int_x^y g(\xi)\,\mathrm{d}\xi \\ > \int_x^y (g(x) - \eta)\,\mathrm{d}\xi = (g(x) - \eta)(y - x), \end{aligned}$$

or

$$\frac{G(y) - G(x)}{y - x} > g(x) - \eta.$$

This implies that

$$\liminf_{y\downarrow x} \frac{G(y) - G(x)}{y - x} \geq g(x)$$

for every $x \in [a,b)$. Therefore, if $x \in [a,b] \setminus C$ we have

$$\begin{aligned}\liminf_{y \downarrow x} &\frac{(G(y) - F(y)) - (G(x) - F(x))}{y - x} \\ &= \liminf_{y \downarrow x} \frac{G(y) - G(x)}{y - x} - \liminf_{y \downarrow x} \frac{F(y) - F(x)}{y - x} \geq g(x) - f(x) > 0.\end{aligned}$$

This implies that, if $x \in [a,b] \setminus C$, there exists $r_x \in \mathbb{R}_{>0}$ such that

$$\frac{(G(y) - F(y)) - (G(x) - F(x))}{y - x} > 0$$

for $y \in (x, x + r_x)$. Since $y - x > 0$ for $y \in (x, x + r_x)$ this implies that

$$(G(y) - F(y)) - (G(x) - F(x)) > 0, \qquad y \in (x, x + r_x).$$

By Lemma 2 this implies that $G - F$ is nondecreasing. Therefore, since $G(a) = F(a)$ it follows that $F(x) \leq G(x)$ for $x \in [a,b]$. Therefore,

$$\begin{aligned}F(x) \leq G(x) &= F(a) + \int_a^x g(\xi)\,\mathrm{d}\xi \\ &= F(a) + \int_a^x f(\xi)\,\mathrm{d}\xi + \int_a^x (g(\xi) - f(\xi))\,\mathrm{d}\xi \\ &\leq F(a) + \int_a^x f(\xi)\,\mathrm{d}\xi + \epsilon\end{aligned}$$

by the definition of g. Since $\epsilon \in \mathbb{R}_{>0}$ is arbitrary, this shows that

$$F(x) \leq F(a) + \int_a^x f(\xi)\,\mathrm{d}\xi.$$

A similar argument to the above, applied to $-F$, gives

$$-F(x) \leq -F(a) - \int_a^x f(\xi)\,\mathrm{d}\xi \quad \Longrightarrow \quad F(x) \geq F(a) + \int_a^x f(\xi)\,\mathrm{d}\xi,$$

which gives the theorem. ■

Our definition of absolute continuity allows us to state a more powerful version of the integration by parts formula than was given as Proposition 3.4.28 for the Riemann integral.

5.9.34 Proposition (Integration by parts) *If* $\mathrm{f}, \mathrm{g}\colon [\mathrm{a}, \mathrm{b}] \to \mathbb{R}$ *are absolutely continuous, then*

$$\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}(\mathrm{x})\mathrm{g}'(\mathrm{x})\,\mathrm{d}\mathrm{x} = \mathrm{f}(\mathrm{b})\mathrm{g}(\mathrm{b}) - \mathrm{f}(\mathrm{a})\mathrm{g}(\mathrm{a}) - \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f}'(\mathrm{x})\mathrm{g}(\mathrm{x})\,\mathrm{d}\mathrm{x}.$$

Proof We have

$$\begin{aligned}& f(x) = f(a) + \int_a^x f'(\xi)\,\mathrm{d}\xi \\ \Longrightarrow\quad & \int_a^b f(x)g'(x)\,\mathrm{d}x = \int_a^b f(a)g'(x)\,\mathrm{d}x + \int_a^b g'(x)\Big(\int_a^x f'(\xi)\,\mathrm{d}\xi\Big)\,\mathrm{d}x \\ \Longrightarrow\quad & \int_a^b f(x)g'(x)\,\mathrm{d}x = f(a)(g(b) - g(a)) + \int_a^b g'(x)\Big(\int_a^b \chi_{[a,x]}(\xi) f'(\xi)\,\mathrm{d}\xi\Big)\,\mathrm{d}x. \qquad (5.22)\end{aligned}$$

By Corollary 5.8.8 the function $F(x,\xi) = g(x)\chi_{[a,x]}(\xi)f'(\xi)$ is integrable with respect to $\lambda_{[a,b]} \times \lambda_{[a,b]}$. Thus we may apply Fubini's Theorem (the version in Theorem 5.8.4) to the last of the above integrals to get

$$\begin{aligned}
\int_a^b g'(x)\Big(\int_a^b \chi_{[a,x]}(\xi)f'(\xi)\,\mathrm{d}\xi\Big)\,\mathrm{d}x &= \int_a^b f'(\xi)\Big(\int_a^b \chi_{[\xi,b]}(x)g'(x)\,\mathrm{d}x\Big)\,\mathrm{d}\xi \\
&= \int_a^b f'(\xi)\Big(\int_\xi^b g'(x)\,\mathrm{d}x\Big)\,\mathrm{d}\xi \\
&= \int_a^b f'(\xi)(g(b)-g(\xi))\,\mathrm{d}\xi \\
&= f(b)g(b) - f(a)g(b) - \int_a^b f'(\xi)g(\xi)\,\mathrm{d}\xi,
\end{aligned}$$

using the fact that $\chi_{[a,x]}(\xi) = \chi_{[\xi,b]}(x)$. Combining this with (5.22) gives the result. ■

5.9.8 Lebesgue points

One might speculate that Lebesgue measurable functions are very nasty. However, in Theorems 5.9.2 and 5.9.3 we show that measurable functions can be approximated well by "nice" functions. In this section we show that if a function is additionally integrable, then we can make some further conclusions about how nice it is.

The main result we state relies on taking a limit over intervals where the length of the interval goes to zero. To make this precise we need to define a directed set for the limit to be well-defined. We refer to ***missing stuff*** for this notion of convergence using directed sets and nets. We let $I \subseteq \mathbb{R}$ be an interval, let $x_0 \in I$, and let $\mathscr{C}(x_0, I)$ be the set of closed subintervals of I containing x_0. This set is partially ordered by saying that $J_1 \preceq J_2$ if $J_1 \supseteq J_2$. It is easily verified that $\mathscr{C}(x_0)$ is a directed set with this partial order. If $f \in \mathsf{L}^{(1)}_{\mathrm{loc}}(I;\mathbb{R})$ then we define $P_{f,x_0} \colon \mathscr{C}(x_0, I) \to \mathbb{R}$ by

$$P_{f,x_0}(J) = \frac{1}{\lambda(J)}\int_J |f(x) - f(x_0)|\,\mathrm{d}x,$$

which defines a $\mathscr{C}(x_0, I)$ net.

5.9.35 Theorem (Almost every point is a Lebesgue point for an integrable function) *If* $\mathrm{f} \in \mathsf{L}^{(1)}(\mathrm{I};\mathbb{R})$ *then* $\lim \mathrm{P}_{\mathrm{f},\mathrm{x}_0} = 0$ *for almost every* $\mathrm{x}_0 \in \mathbb{R}$.

Proof We begin with a technical lemma.

1 Lemma *If* $\mathrm{f} \in \mathsf{L}^{(1)}([\mathrm{a},\mathrm{b}];\mathbb{R})$ *then there exists* $\mathrm{A} \subseteq [\mathrm{a},\mathrm{b}]$ *such that*

(i) $\lambda([\mathrm{a},\mathrm{b}] \setminus \mathrm{A}) = 0$ *and such that*

(ii) for all $\alpha \in \mathbb{R}$ *and for all* $\mathrm{x} \in \mathrm{A}$,

$$\lim_{\delta\downarrow 0}\frac{1}{\delta}\int_{\mathrm{x}}^{\mathrm{x}+\delta}|\mathrm{f}(\xi)-\alpha|\,\mathrm{d}\xi = \lim_{\delta\downarrow 0}\frac{1}{\delta}\int_{\mathrm{x}-\delta}^{\mathrm{x}}|\mathrm{f}(\xi)-\alpha|\,\mathrm{d}\xi = |\mathrm{f}(\mathrm{x})-\alpha|.$$

Proof Let $\alpha \in \mathbb{R}$, let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be an enumeration of the rationals and, for $j \in \mathbb{Z}_{>0}$, define $f_j \in \mathsf{L}^{(1)}([a,b];\mathbb{R})$ by

$$f_j(x) = |f(x) - \alpha|.$$

By part (ii) of Theorem 5.9.31, for each $j \in \mathbb{Z}_{>0}$ there exists a set $A_j \subseteq [a,b]$ such that $\lambda([a,b] \setminus A_j) = 0$ and such that, for all $x \in A_j$,

$$\lim_{\delta\downarrow 0} \frac{1}{\delta}\int_x^{x+\delta} g_j(\xi)\,\mathrm{d}\xi = \lim_{\delta\downarrow 0} \frac{1}{\delta}\int_{x-\delta}^{x} g_j(\xi)\,\mathrm{d}\xi = g_j(x).$$

Take $A = \cap_{j\in\mathbb{Z}_{>0}} A_j$, and note that

$$\lambda([a,b] \setminus A) = \lambda(\cup_{j\in\mathbb{Z}_{>0}}[a,b] \setminus A_j) = 0,$$

where we have used De Morgan's Laws and Exercise 2.5.9.

Let $\delta \in \mathbb{R}_{>0}$ and let $k \in \mathbb{Z}_{>0}$ be such that $|q_k - \alpha| < \frac{\delta}{3}$. By Exercise 2.2.7 we have

$$||f(x) - \alpha| - |f(x) - q_k|| \le |q_j - \alpha| < \tfrac{\delta}{3}$$

for all $x \in [a,b]$. Therefore,

$$\left|\frac{1}{\delta}\int_x^{x+\delta} |f(\xi) - \alpha|\,\mathrm{d}\xi - \frac{1}{\delta}\int_x^{x+\delta} |g_k(\xi) - \alpha|\,\mathrm{d}\xi\right| \le frac1\delta \int_x^{x+\delta} \frac{\delta}{3}\,\mathrm{d}\xi = \frac{\delta}{3}$$

for every $\delta \in \mathbb{R}_{>0}$ such that the integrals are defined. Therefore, we let $x \in A$ and let $\delta_0 \in \mathbb{R}_{>0}$ be such that

$$\left|\frac{1}{\delta}\int_x^{x+\delta} g_k(\xi)\,\mathrm{d}\xi - g_k(x)\right| < \frac{\delta}{3}$$

for all $\delta \in (0,\delta_0)$. Then, provided that $\delta \in (0,\delta_0)$ we have

$$\begin{aligned}\left|\frac{1}{\delta}\int_x^{x+\delta} |f(\xi) - \alpha|\,\mathrm{d}\xi - f(x)\right| \le{}& \left|\frac{1}{\delta}\int_x^{x+\delta} |f(\xi) - \alpha|\,\mathrm{d}\xi - \frac{1}{\delta}\int_x^{x+\delta} |g_k(\xi) - \alpha|\,\mathrm{d}\xi\right| \\ &+ \left|\frac{1}{\delta}\int_x^{x+\delta} g_k(\xi)\,\mathrm{d}\xi - g_k(x)\right| + |q_k - \alpha| < \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta,\end{aligned}$$

using the triangle inequality. This gives the left limit equal to the right expression in the statement of the lemma. The proof that the middle limit is equal to the right expression follows along entirely similar lines. ▼

It is now somewhat easy to complete the proof of the theorem. First suppose that $I = [a,b]$ is compact. As per the preceding lemma, let $A \subseteq [a,b]$ be such that $\lambda([a,b] \setminus A) = 0$ and such that, for all $\alpha \in \mathbb{R}$ and $x \in A$,

$$\lim_{\delta\downarrow 0} \frac{1}{\delta}\int_x^{x+\delta} |f(\xi) - \alpha|\,\mathrm{d}\xi = \lim_{\delta\downarrow 0} \frac{1}{\delta}\int_{x-\delta}^{x} |f(\xi) - \alpha|\,\mathrm{d}\xi = |f(x) - \alpha|.$$

Now let $x_0 \in A \cap (a,b)$ and let $\epsilon \in \mathbb{R}_{>0}$. Then there exists $\delta_0 \in \mathbb{R}_{>0}$ such that

$$\left|\frac{1}{\delta}\int_{x_0}^{x_0+\delta} |f(x) - f(x_0)|\,\mathrm{d}x\right| < \frac{\epsilon}{2}$$

and

$$\Big|\frac{1}{\delta}\int_{x_0-\delta}^{x_0}|f(x)-f(x_0)|\,\mathrm{d}x\Big| < \frac{\epsilon}{2}$$

for $\delta \in (0,\delta_0)$. Define $J_0 = [x_0-\delta_0, x_0+\delta_0]$. We may suppose that δ_0 is sufficiently small that $J_0 \in [a,b]$. If $J_0 \preceq J$ then $J \subseteq J_0$ and so we have

$$\begin{aligned}\Big|\frac{1}{\lambda(J)}\int_J|f(x)-f(x_0)|\,\mathrm{d}x\Big|& \\ \leq \Big|\frac{1}{\delta_0}\int_{x_0}^{x_0+\delta_0}&|f(x)-f(x_0)|\,\mathrm{d}x + \frac{1}{\delta_0}\int_{x_0-\delta_0}^{x_0}|f(x)-f(x_0)|\,\mathrm{d}x\Big| \leq \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon.\end{aligned}$$

This shows that $\lim P_{f,x_0} = 0$, giving the theorem when I is compact. When I is not compact, then we can write I as a countable union of compact intervals $(I_j)_{j\in\mathbb{Z}_{>0}}$. For each $j \in \mathbb{Z}_{>0}$ let $N_j \subseteq I_j$ be the set of measure zero such that if $x \in N_j$ we have $\lim P_{f,x} \neq 0$. Since $\lambda(\cup_{j\in\mathbb{Z}_{>0}}N_j) = 0$ by Exercise 2.5.9, and since if $x \in I \setminus \cup_{j\in\mathbb{Z}_{>0}}N_j$ we have $\lim P_{f,x} = 0$, the theorem follows. ■

5.9.9 Maximal functions

5.9.10 The change of variables formula

In this section we state and prove a simple version of the change of variables formula for the Lebesgue integral. This is one of the places in our development where the extension to the Lebesgue integral on $\mathbb{R}^n$ is not so easily accomplished. Indeed, the higher-dimensional versions are difficult to prove in any useful degree of generality, and normally require the Radon–Nikodym Theorem (see, for example, **WR:86**). We refer to **[DEV:71]** for a quite general statement of the multivariable change of variable formula. Fortunately, we shall only need the single-variable change of variable, and this can be proved more directly, even though, as the reader can see, the proof is not quite trivial.

5.9.36 Theorem (Change of variable) *Let* $\mathrm{I}, \mathrm{J} \subseteq \mathbb{R}$ *be intervals with* $\phi\colon \mathrm{I} \to \mathrm{J}$ *a map with the properties that*

(i) *ϕ is surjective,*

(ii) *ϕ is either monotonically decreasing or monotonically increasing, and*

(iii) *there exists an integrable function* $\phi'\colon \mathrm{I} \to \mathbb{R}$ *and* $\mathrm{x}_0 \in \mathrm{I}$ *so that**missing stuff***

$$\phi(\mathrm{x}) = \phi(\mathrm{x}_0) + \int_{[\mathrm{x}_0,\mathrm{x}]}\phi'\,\mathrm{d}\lambda_{[\mathrm{x}_0,\mathrm{x}]}.$$

If $\mathrm{f}\colon \mathrm{J} \to \mathbb{R}$ *is integrable then* $\mathrm{f}\circ\phi$ *is measurable,* $\mathrm{f}\circ\phi|\phi'|$ *is integrable, and*

$$\int_{\mathrm{J}}\mathrm{f}\,\mathrm{d}\lambda_{\mathrm{J}} = \int_{\mathrm{I}}\mathrm{f}\circ\phi|\phi'|\,\mathrm{d}\lambda_{\mathrm{I}}.$$

Proof We first take the case where $I = [a,b]$ and $J = [c,d]$. We claim that the theorem is true for step functions in this case. Indeed, let $g\colon [c,d] \to \mathbb{R}$ be a step function and write

$$g = \sum_{j=1}^{k} \alpha_j \chi_{I_j}$$

where $I_j = (x_j, x_{j-1}]$, $j \in \{0, 1, \dots, k\}$, are the endpoints of a partition $(I_1, \dots, I_k)$ of $[c, d]$. Corresponding to this partition of $[c, d]$ we define a partition $(J_1, \dots, J_k)$ of $[a, b]$ endpoints $(\xi_0, \xi_1, \dots, \xi_k)$ such that $\phi(\xi_j) = x_j$, $j \in \{0, 1, \dots, k\}$. There may be ambiguity in this definition of ξ_j, $j \in \{0, 1, \dots, k\}$, but this does not matter. Assuming that $\phi'(x) \geq 0$ for all x we then compute

$$\begin{aligned}\int_a^b g \circ \phi(\xi) \phi'(\xi)\, \mathrm{d}\xi &= \sum_{j=1}^{k} \int_{\xi_{j-1}}^{\xi_j} \alpha_j \phi'(\xi)\, \mathrm{d}\xi \\ &= \sum_{j=1}^{k} \alpha_j (\phi(\xi_j) - \phi(\xi_{j-1})) \\ &= \sum_{j=1}^{k} \alpha_j (x_j - x_{j-1}) = \int_c^d g(x)\, \mathrm{d}x.\end{aligned}$$

A similarly styled computation shows that the result is also true if $\phi'(x) \leq 0$.

Now suppose that f takes values in $[0, \infty)$. Using Theorem 5.9.2, let $(g_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence of step functions on $[c, d]$ with the property that for almost every $x \in [c, d]$ we have $\lim_{j \to \infty} g_j(x) = f(x)$. Let us denote by Z_1 the subset of measure zero where this limit does not hold. By examining the proofs of Proposition 5.6.39 and Theorem 5.9.2 we see that we can take the sequence $(g_j)_{j \in \mathbb{Z}_{>0}}$ so that for each $x \in [c, d] \setminus Z_1$ the sequence $(g_j(x))_{j \in \mathbb{Z}_{>0}}$ is nondecreasing. Therefore the sequence $(g_j \circ \phi(\xi) |\phi'(\xi)|)_{j \in \mathbb{Z}_{>0}}$ is also nondecreasing provided that $\phi(\xi) \notin Z_1$. Indeed, provided that either

1. $\phi(\xi) \notin Z_1$ or
2. $\phi'(x) = 0$

hold, then we have $\lim_{j \to \infty} g_j \circ \phi(\xi) |\phi'(\xi)| = f \circ \phi(\xi) |\phi'(\xi)|$. We claim that the set of points $Z_2 \subseteq [a, b]$ where both conditions 1 and 2 fail to hold has measure zero. We do this with the aid of a lemma.

1 Lemma $\mathrm{Z} \subseteq \mathbb{R}$ *has Lebesgue measure zero if and only if there exists a sequence* $(\mathrm{I_j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ *of nonempty open intervals such that*

(i) $\sum_{\mathrm{j}=1}^{\infty} \lambda(\mathrm{I_j}) < \infty$ *and*

(ii) for each $\mathrm{x} \in \mathrm{Z}$ *there exists a sequence* $(\mathrm{j_k})_{\mathrm{k} \in \mathbb{Z}_{>0}}$ *of* $\mathbb{Z}_{>0}$ *so that* $\mathrm{x} \in \mathrm{I_{j_k}}$, $\mathrm{k} \in \mathbb{Z}_{>0}$.

Proof By definition, Z has Lebesgue measure zero if for each $\epsilon \in \mathbb{R}_{>0}$ there exists a family $(\tilde{I}_\ell)_{\ell \in \mathbb{Z}_{>0}}$ of open intervals, some possibly empty, for which $Z \subseteq \cup_{\ell \in \mathbb{Z}_{>0}} \tilde{I}_\ell$ and $\sum_{\ell=1}^{\infty} \lambda(\tilde{I}_\ell) < \epsilon$.

Suppose that there exists a collection of intervals $(I_j)_{j \in \mathbb{Z}_{>0}}$ having properties (i) and (ii) and let $\epsilon \in \mathbb{R}_{>0}$. Choose a finite collection $I_{j_1}, \dots, I_{j_m}$ of intervals so that

$$\sum_{j=1}^{\infty} \lambda(I_j) - \sum_{k=1}^{m} \lambda(I_{j_k}) < \epsilon.$$

It then follows that the family $(I_j)_{j\in\mathbb{Z}_{>0}} \setminus (I_{j_k})_{k\in\{1,\dots,m\}}$ of open intervals has total length less than ϵ. Furthermore, since only a finite number of intervals are removed from $(I_j)_{j\in\mathbb{Z}_{>0}}$, the remaining intervals still cover Z. Thus Z has Lebesgue measure zero.

Now suppose that Z has measure zero. For $n \in \mathbb{Z}_{>0}$ let $(I_{n,j})_{j\in\mathbb{Z}_{>0}}$ have the property that $Z \subseteq \cup_{j\in\mathbb{Z}_{>0}} I_{n,j}$ and that

$$\sum_{j=1}^{\infty} \lambda(I_{n,j}) < \frac{1}{2^n}.$$

Then the collection $(I_{j,n})_{j,n\in\mathbb{Z}_{>0}}$ satisfies (i) and (ii). ▼

According to the lemma, choose a sequence $(I_j)_{j\in\mathbb{Z}_{>0}}$ of intervals covering Z_1 and whose total length is finite. Define a step function $g_n\colon [c,d] \to \mathbb{R}$ by

$$g_n = \sum_{j=1}^{n} \chi_{I_j},$$

and note that for each $x \in Z_1$ we have $\lim_{n\to\infty} g_n(x) = \infty$. If $\xi \in Z_2$ it follows that $\lim_{n\to\infty} g_n \circ \phi(\xi)|\phi'(\xi)| = \infty$. Now note that

$$\int_a^b g_n \circ \phi(\xi)|\phi'(\xi)|\,\mathrm{d}\xi = \int_c^d g_n(x)\,\mathrm{d}x < \sum_{j=1}^{\infty} \lambda(I_j) < \infty.$$

It follows from Proposition 5.7.12 that $\lambda(Z_2) = 0$.

Thus we have shown that, provided that $I = [a,b]$, that $J = [c,d]$, and that $f(J) \subseteq [0,\infty)$, for almost every $\xi \in [a,b]$ and almost every $x \in [c,d]$ we have

$$\lim_{j\to\infty} g_j(x) = f(x), \quad \lim_{j\to\infty} g_j \circ \phi(\xi)|\phi'(\xi)| = f \circ \phi(\xi)|\phi'(\xi)|$$

with both limits being monotonic, and that for each $j \in \mathbb{Z}_{>0}$ we have

$$\int_a^b g_j \circ \phi(\xi)\,\mathrm{d}\xi = \int_c^d g_j(x)\,\mathrm{d}x.$$

The result under the current assumptions now follows by the Monotone Convergence Theorem. For an arbitrary f with $I = [a,b]$ and $J = [c,d]$ the result follows from breaking f into its positive and negative parts.

It remains to prove the result for general intervals I and J. Let $(I_n = [a_n, b_n])_{j\in\mathbb{Z}_{>0}}$ be a sequence of intervals with the property that $\operatorname{int}(I) = \cup_{n\in\mathbb{Z}_{>0}} I_n$. Define $J_n = \phi(I_n)$, $n \in \mathbb{Z}_{>0}$, noting that J_n so defined is a closed interval by monotonicity of ϕ. We then have, by the Dominated Convergence Theorem,

$$\int_I f\,\mathrm{d}\lambda_I = \lim_{n\to\infty} \int_I \chi_{I_n} f\,\mathrm{d}\lambda_I, \quad \int_J f \circ \phi|\phi'|\,\mathrm{d}\lambda_J = \lim_{n\to\infty} \int_J \chi_{J_n} f \circ \phi|\phi'|\,\mathrm{d}\lambda_J.$$

From this the result follows since

$$\int_I \chi_{I_n} f\,\mathrm{d}\lambda_I = \int_J \chi_{J_n} f \circ \phi|\phi'|\,\mathrm{d}\lambda_J.$$

■

5.9.11 Topological characterisations of the deficiencies of the Riemann integral[16]

In Section 5.7.5 we saw that it was possible to give interesting topological characterisations of the Dominated Convergence Theorem for the general measure theoretic integral. These characterisations are, of course, inherited by the Lebesgue integral. That is to say, one can specialise Theorems 5.7.40 and 5.7.42 to the Lebesgue integral as follows.

5.9.37 Theorem (Topological "everywhere" Dominated Convergence Theorem for the Lebesgue integral) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *then* C_p*-bounded subsets of* $\mathsf{L}^{(1)}(\mathrm{A};\mathbb{R})$ *are* C_p*-sequentially closed.*

5.9.38 Theorem (Limit structure "almost everywhere" Dominated Convergence Theorem for the Lebesgue integral) *If* $\mathrm{A} \in \mathscr{L}(\mathbb{R})$ *then* $\mathscr{L}_{\lambda_\mathrm{A}}$*-bounded subsets of* $\mathsf{L}^1(\mathrm{A};\mathbb{R})$ *are* $\mathscr{L}_{\lambda_\mathrm{A}}$*-sequentially closed.*

In this section we give a couple of examples that show that these theorems do not hold for the Riemann integral. First we consider the "everywhere" version of the Dominated Convergence Theorem.

5.9.39 Example (The topological "everywhere" Dominated Convergence Theorem does not hold for the Riemann integral) By means of an example, we show that there are C_p-bounded subsets of the seminormed vector space $\mathsf{R}^{(1)}([0,1];\mathbb{R})$ that are not C_p-sequentially closed. Let us denote

$$B = \{f \in \mathsf{R}^{(1)}([0,1];\mathbb{R}) \mid \ |f(x)| \leq 1\},$$

noting by Proposition 5.7.39 that B is C_p-bounded. Let $(q_j)_{j\in\mathbb{Z}_{>0}}$ be an enumeration of the rational numbers in $[0,1]$ and define a sequence $(f_k)_{k\in\mathbb{Z}_{>0}}$ in $\mathsf{R}^{(1)}([0,1];\mathbb{R})$ by

$$f_k(x) = \begin{cases} 1, & x \in \{q_1,\dots,q_k\}, \\ 0, & \text{otherwise.} \end{cases}$$

The sequence converges in the C_p-topology to the characteristic function of $\mathbb{Q}\cap[0,1]$; let us denote this function by f. This limit function is not Riemann integrable and so not in $\mathsf{R}^{(1)}([0,1];\mathbb{R})$. Thus B is not C_p-sequentially closed. •

Next we turn to the "almost everywhere" version of the Dominated Convergence Theorem for the Riemann integral.

5.9.40 Example (The limit structure "almost everywhere" Dominated Convergence Theorem does not hold for the Riemann integral) Recall from Section 5.6.6 that $\mathsf{L}^0([0,1];\mathbb{R})$ denotes the set of equivalence classes of $\mathbb{R}$-valued measurable functions on $[0,1]$ under the equivalence relation of almost everywhere equality. We denote by $\mathsf{R}^1([0,1];\mathbb{R})$ the image of $\mathsf{R}^{(1)}([0,1];\mathbb{R})$ by the projection from $\mathsf{L}^{(0)}([0,1];\mathbb{R})$ to

[16] The results in this section are not used in an essential way anywhere else in the text.

$\mathsf{L}^0([0,1];\mathbb{R})$. Thus elements of $\mathsf{R}^1([0,1];\mathbb{R})$ are equivalence classes of $\mathbb{R}$-valued Riemann integrable functions under the equivalence relation of almost everywhere equality. We denote elements of $\mathsf{R}^1([0,1];\mathbb{R})$ by $[f]$, reflecting the fact that they are equivalence classes of functions. For brevity we denote the Lebesgue measure on $[0,1]$ by λ.

We give an example that shows that $\mathscr{L}_\lambda$-bounded subsets of the normed vector space $\mathsf{R}^1([0,1];\mathbb{R})$ are not $\mathscr{L}_\lambda$-sequentially closed. We first remark that the construction of Example 5.9.39, projected to $\mathsf{R}^1([0,1];\mathbb{R})$, does not suffice because $[f]$ is equal to the equivalence class of the zero function which *is* Riemann integrable, even though f is not. The fact that $[f]$ contains functions that are Riemann integrable and functions that are not Riemann integrable is a reflection of the fact that the set

$$\mathsf{R}_0([0,1];\mathbb{R}) = \left\{f\colon [0,1] \to \mathbb{R} \;\middle|\; f \text{ Riemann integrable and } \int_0^1 f(x)\,\mathrm{d}x = 0\right\}$$

is not sequentially closed. This is a phenomenon of interest, but it is not what is of interest here.

We use the construction of the function f from the proof of Proposition 5.1.12. In that proof, the function f was shown to have the following properties:

1. f is the pointwise limit of a sequence $(f_k)_{k\in\mathbb{Z}_{>0}}$ of Riemann integrable functions;
2. any function almost everywhere equal to f is not Riemann integrable.

Therefore, by Theorem 5.6.51 it follows that $([f_k])_{k\in\mathbb{Z}_{>0}}$ is $\mathscr{L}_\lambda$-convergent to $[f]$. Moreover, $[f] \notin \mathsf{R}^1([0,1];\mathbb{R})$. To complete the example, we note that the sequence $([G_k])_{j\in\mathbb{Z}_{>0}}$ is in the set

$$B = \{[f] \in \mathsf{R}^1([0,1];\mathbb{R}) \mid |f(x)| \leq 1 \text{ for almost every } x \in [0,1]\},$$

which is $\mathscr{L}_\lambda$-bounded by Proposition 5.7.41. The example shows that this $\mathscr{L}_\lambda$-bounded subset of $\mathsf{R}^1([0,1];\mathbb{R})$ is not $\mathscr{L}_\lambda$-sequentially closed. •

Exercises

5.9.1 Use Lemma 5.9.30 to directly conclude that the Cantor function of Example 5.9.25 is not absolutely continuous.

5.9.2 Give an example of a function $f\colon \mathbb{R} \to \mathbb{R}$ such that $|f|$ is Lebesgue measurable, but f is not Lebesgue measurable.

5.9.3 Answer the following two questions.

(a) Why must a Riemann integrable function $f\colon [a,b] \to \mathbb{R}$ on a compact interval be bounded?

(b) Provide an unbounded function on $[a,b]$ that is continuous when restricted to (a,b), and that is Lebesgue integrable.

One of the differences between the Lebesgue and Riemann integral is that the Lebesgue integral is defined by first approximating a measurable function by a sequence of simple function only from below. In contrast, for the Riemann integral, one asks that the function be approximated from below *and* above by step functions. One might legitimately wonder whether this is asking too much of the approximation, and whether one can get away, as one does with the Lebesgue integral, by approximation from (say) below. The following exercise asks you to explore this.

5.9.4 Let $I = [0,1]$ and let

$$f = \chi_{I\cap\mathbb{Q}}, \quad g = \chi_{I\cap(\mathbb{R}\setminus\mathbb{Q})}.$$

Answer the following questions.

(a) Show that $I_-(f) = I_-(g) = 0$. Thus, when approximated just by step functions from below, both f and g have zero "integral."

(b) Show that $I_-(f+g) \neq I_-(f) + I_-(g)$. Thus the "integral" is not linear.

5.9.5 Let $A \subseteq I = [0,1]$ be the subset of irrational numbers, and let χ_A be the characteristic function. Show that $\int_I \chi_A \, \mathrm{d}\lambda = 1$.

5.9.6 Show that there is a function $f\colon [0,1] \to \mathbb{R}$ that is not Riemann integrable, but for which $|f|$ is Riemann integrable.

5.9.7 Let $I = [0,\infty)$ and define $f\colon I \to \mathbb{R}$ by $f(x) = x$. Use the Monotone Convergence Theorem to show that f is not integrable.

5.9.8 Let $I \subseteq \mathbb{R}$ be an interval, and let $f\colon I \to \mathbb{R}$ be continuous. Show that if

$$\lambda(\{x \in I \mid f(x) \neq 0\}) = 0$$

then $f(x) = 0$ for every $x \in I$.

Section 5.10

Notes

Chapter 6

Banach spaces

In Chapter **??**, particularly in Sections **??** and **??**, we studied linear algebra over arbitrary fields. Here we relied on the notion, introduced in Section 4.3, of a vector space. In many instances in applications, one is interested in the case where the field is either $\mathbb{R}$ or $\mathbb{C}$. In finite-dimensions, the story here is not too complicated; finite-dimensional vector spaces over $\mathbb{R}$ or $\mathbb{C}$ are fairly easy to understand and linear maps on these spaces are also fairly easy to understand. However, in applications, it turns out that infinite-dimensional vector spaces are often what is of most interest. We make no attempt to motivate this here, but refer the reader to Chapter 8. The reader will note that we were careful to understand the algebra of infinite-dimensional vector spaces in Section 4.3 and linear maps between them in Section **??**. It turns out, though, that the key to understanding the infinite-dimensional vector spaces that arise in applications is through the various topologies one can put on these. This is the genesis of the huge subject of topological vector spaces which we spend the next three chapters introducing. The present chapter is devoted to topologies defined by a "norm." These are the most basic topologies, and suffice to cover many, but by no means all, areas of application.

Certain parts of what we say in this chapter have already been accounted for in Chapter **??**. However, we it seems like a good idea to make the treatment here independent, for the most part, of the more general and abstract treatment in Chapter **??**. Therefore, at the cost of repetitiveness we make treat all of the topological ideas for normed vector spaces independently of the fact that we have already considered them.

Do I need to read this chapter? This chapter is fundamental to understanding in any rigorous way topics like Fourier series, Fourier transforms, linear system theory, signal processing, etc. This makes at least the basic material in this chapter essential reading. Perhaps a reading of the detailed examples of dual spaces in Section **??** can be postponed until it is needed, although it is at least interesting. •

Contents

6.1 Definitions and properties of normed vector spaces 604
6.1.1 Norms and seminorms 604
6.1.2 Open and closed subsets of normed vector spaces 610

6.1.3 Subspaces, direct sums, and quotients 614
6.2 Sequences in normed vector spaces 619
6.2.1 Definitions and properties of sequences 619
6.2.2 Algebraic operations on sequences 621
6.2.3 Multiple sequences 621
6.3 Completeness and completions 623
6.3.1 Completeness (Banach spaces) 623
6.3.2 Why completeness is important 627
6.3.3 Completeness and direct sums and quotients 629
6.3.4 Completions 630
6.4 Series in normed vector spaces 634
6.4.1 Definitions and properties of series 634
6.4.2 Absolute and unconditional convergence 636
6.4.3 Algebraic operations on series 645
6.4.4 Multiple series 647
6.4.5 Cesàro convergence of sequences and series 648
6.4.6 Series in normed vector spaces with arbitrary index sets 650
6.4.7 Notes 655
6.5 Continuous maps between normed vector spaces 656
6.5.1 General continuous maps between normed vector spaces 656
6.5.2 Continuous linear maps between normed vector spaces 659
6.5.3 Induced topologies on continuous linear maps 664
6.5.4 The Open Mapping Theorem and Closed Graph Theorem 670
6.6 Topology of normed vector spaces 672
6.6.1 Properties of balls in normed vector spaces 672
6.6.2 Interior, closure, boundary, etc. 673
6.6.3 Compactness 678
6.6.4 Closed subspaces 680
6.6.5 Bases for normed vector spaces 689
6.6.6 Notes 692
6.7 Examples of Banach spaces 693
6.7.1 The p-norms on $\mathbb{F}^n$ 693
6.7.2 Banach spaces of sequences 700
6.7.3 Banach spaces of direct sums 709
6.7.4 Banach spaces of continuous functions on $\mathbb{R}$ 713
6.7.5 Banach spaces of continuous functions on metric spaces 720
6.7.6 Banach spaces of continuous functions on locally compact topological spaces 721
6.7.7 Banach spaces of integrable functions on $\mathbb{R}$ 721
6.7.8 Banach spaces of integrable functions on measure spaces 732
6.7.9 Banach spaces of measures 732
6.7.10 Notes 734

Section 6.1

Definitions and properties of normed vector spaces

The basic ingredient in this chapter is a norm on a vector space. While it is possible to introduce this notion for other classes of fields, we restrict our attention to vector spaces over either $\mathbb{R}$ or $\mathbb{C}$. It will often be convenient to be able to consider both of these cases together, and so let us introduce some notation for doing this.

6.1.1 Notation ($\mathbb{F}$) The symbol $\mathbb{F}$ will denote either $\mathbb{R}$ or $\mathbb{C}$. That is to say, whenever the symbol $\mathbb{F}$ is present, the statement can be read by replacing it with either $\mathbb{R}$ or $\mathbb{C}$. In order to use this convenient notation as much as possible we have the following conventions.

(i) If $\mathbb{F} = \mathbb{R}$ and if $a \in \mathbb{F}$ then $|a|$ denotes the absolute value of a.

(ii) If $\mathbb{F} = \mathbb{C}$ and if $a \in \mathbb{F}$ then $|a|$ denotes the modulus of a.

(iii) If $\mathbb{F} = \mathbb{R}$ and if $a \in \mathbb{F}$ then $\bar{a} = a$.

(iv) If $\mathbb{F} = \mathbb{C}$ if $a \in \mathbb{F}$ then $\bar{a}$ is the complex conjugate of a. •

Do I need to read this section? Accepting that normed vector spaces are important (they are), this section must then be important. •

6.1.1 Norms and seminorms

In this section we consider norms and seminorms. While the notion of a norm is the most important for us, we will see that seminorms come up in two natural ways. One is in Section 6.7.8 when we give an extremely important class of normed vector spaces. As we shall see, in the construction of this class it is natural to first define a seminorm. Thus, although one is interested in a norm in the end, a seminorm naturally arises along the way. In a completely different manner, seminorms will be important in Chapter **??** in their own right. As we shall see, particularly in the context of so-called "generalised signals" in Chapter 10, seminorms often arise in natural way independently of whether they are used to define a norm.

In any event, here are the definitions.

6.1.2 Definition (Seminorm, norm) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be an $\mathbb{F}$-vector space. A ***seminorm*** on V is a map $\mathsf{V} \ni v \mapsto \|v\| \in \mathbb{R}_{\geq 0}$ with the following properties:

(i) $\|av\| = |a|\|v\|$ for $a \in \mathbb{F}$ and $v \in \mathsf{V}$ (***homogeneity***);

(ii) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$ for $v_1, v_2 \in \mathsf{V}$ (***triangle inequality***).

A ***norm*** on V is a seminorm $v \mapsto \|v\|$ with the additional property that

(iii) $\|v\| = 0$ only if $v = 0_{\mathsf{V}}$ (***positive-definiteness***).

We shall often denote a seminorm by $\|\cdot\|$. •

Let us give some examples of norms and seminorms. Sometimes examples are illustrative and sometimes they are of great value in their own right. The examples below, with the exception of the first one, are all of great independent interest, as well as illustrating the concept of a norm.

6.1.3 Examples (Seminorm, norm)

1. For any $\mathbb{F}$-vector space V there is a useless seminorm defined by $v \mapsto 0$. Let us call this the ***trivial seminorm*** since it is good for giving trivial examples. Unless $\mathsf{V} = \{0_{\mathsf{V}}\}$, the trivial seminorm is never a norm.
2. On $\mathbb{F}^n$ define
$$\|\boldsymbol{v}\|_2 = \left(|v_1|^2 + \cdots + |v_n|^2\right)^{1/2}.$$
In the case when $\mathbb{F} = \mathbb{R}$ this is the standard norm on $\mathbb{R}^n$ as discussed in Section **??**. In particular, this norm defines the usual notion of length of a vector in $\mathbb{F}^n$, i.e., $\|\boldsymbol{v}\|$ is the distance from $\boldsymbol{0}_{\mathbb{F}^n}$ to $\boldsymbol{v}$. Note that we now use different notation for this norm. We shall also sometimes call it the **2-*norm*** on $\mathbb{F}^n$ rather than the standard norm. It is pretty evident that $\|\cdot\|_2$ satisfies the homogeneity and positive-definiteness properties required of a norm. It is also true that $\|\cdot\|_2$ satisfies the triangle inequality. We do not prove this here, although it was proved in the case when $\mathbb{F} = \mathbb{R}$ as part of Proposition **??**. The proof of this relies on the so-called "Cauchy–Bunyakovsky–Schwarz Inequality." This inequality holds because $\|\cdot\|_2$ is the norm derived from an inner product on $\mathbb{F}^n$. Thus we shall see how $\|\cdot\|_2$ satisfies the triangle inequality when we discuss inner products in Section 7.1. Moreover, we shall see this example come up in another general context in Section 6.7.1. The point is that we will subsequently see multiple proofs of the triangle inequality for $\|\cdot\|_2$.
3. Let us consider another norm on $\mathbb{F}^n$ which differs from the standard norm. For $\boldsymbol{v} = (v_1, \dots, v_n) \in \mathbb{F}^n$ define
$$\|\boldsymbol{v}\|_1 = |v_1| + \cdots + |v_n|.$$
All properties of the norm are readily verified, including the triangle inequality, as this now follows from the triangle inequality for $|\cdot|$. Although different from the standard norm, this norm is in some sense equivalent to it, and we refer to Exercise 6.1.6 for an exploration of this. This norm is called the **1-*norm***.
4. Let us consider a final (for now) norm on $\mathbb{F}^n$ given by
$$\|\boldsymbol{v}\|_\infty = \max\{|v_j| \mid j \in \{1, \dots, n\}\}.$$
This is in fact a norm, called the **∞-*norm***. The only not entirely trivial norm property to verify is the triangle inequality. For this, let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}^n$ and let $j, k, \ell \in \{1, \dots, n\}$ have the property that $\|\boldsymbol{u}\|_\infty = |u_j|$, $\|\boldsymbol{v}\|_\infty = |v_k|$, and $\|\boldsymbol{u} + \boldsymbol{v}\|_\infty = |u_\ell + v_\ell|$. We then have
$$\|\boldsymbol{u} + \boldsymbol{v}\|_\infty = |u_\ell + v_\ell| \le |u_\ell| + |v_\ell| \le |u_j| + |v_k| = \|\boldsymbol{u}\|_\infty + \|\boldsymbol{v}\|_\infty.$$
Note that this norm is also different from the standard norm, but it is equivalent in some sense; Exercise 6.1.6.

The above three examples of norms were all defined on the finite-dimensional $\mathbb{F}$-vector space $\mathbb{F}^n$. Let us now consider infinite-dimensional analogues of these norms.

5. Recall from Example 4.3.2–**??** that $\mathbb{F}_0^\infty$ denotes the sequences $(v_j)_{j\in\mathbb{Z}_{>0}}$ for which the set $\{j \in \mathbb{Z}_{>0} \mid v_j \neq 0\}$ is finite. Thus sequences in $\mathbb{F}_0^\infty$ are eventually zero. We define

   $$\|(v_j)_{j\in\mathbb{Z}_{>0}}\|_2 = \Big(\sum_{j=1}^{\infty}|v_j|^2\Big)^{1/2},$$

   noting that the sum makes sense since it is actually finite. That $\|\cdot\|_2$ satisfies the properties of a norm is straightforward. Let us verify just the triangle inequality, since its proof gives the idea of how the norm works. We let $(u_j)_{j\in\mathbb{Z}_{>0}}, (v_j)_{j\in\mathbb{Z}_{>0}} \in \mathbb{F}_0^\infty$ and let $N \in \mathbb{Z}_{>0}$ be such that $u_j = v_j = 0$ for $j \geq N$. Then

   $$\begin{aligned}\|(u_j)_{j\in\mathbb{Z}_{>0}} + (v_j)_{j\in\mathbb{Z}_{>0}}\|_2 &= \Big(\sum_{j=1}^{\infty}|u_j|^2 + \sum_{j=1}^{\infty}|v_j|^2\Big)^{1/2} = \Big(\sum_{j=1}^{\infty}(|u_j|^2 + |v_j|^2)\Big)^{1/2}\\ &= \Big(\sum_{j=1}^{N}(|u_j|^2 + |v_j|^2)\Big)^{1/2} \leq \Big(\sum_{j=1}^{N}|u_j|^2\Big)^{1/2} + \Big(\sum_{j=1}^{N}|u_j|^2\Big)^{1/2}\\ &= \Big(\sum_{j=1}^{\infty}|u_j|^2\Big)^{1/2} + \Big(\sum_{j=1}^{\infty}|u_j|^2\Big)^{1/2}\\ &= \|(u_j)_{j\in\mathbb{Z}_{>0}}\|_2 + \|(v_j)_{j\in\mathbb{Z}_{>0}}\|_2,\end{aligned}$$

   where we have used the triangle inequality for the 2-norm on $\mathbb{F}^N$. This norm is called the ***2-norm*** on $\mathbb{F}_0^\infty$.
6. We again consider the vector space $\mathbb{F}_0^\infty$ and now define

   $$\|(v_j)_{j\in\mathbb{Z}_{>0}}\|_1 = \sum_{j=1}^{\infty}|v_j|,$$

   this sum again making sense since it is finite. It is easy to verify, just as we did for the 2-norm above, that $\|\cdot\|_1$ is a norm, and we call it the ***1-norm***.
7. As a final norm on $\mathbb{F}_0^\infty$ we define

   $$\|(v_j)_{j\in\mathbb{Z}_{>0}}\|_\infty = \sup\{|v_j| \mid j \in \mathbb{Z}_{>0}\}.$$

   Because the sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ is finite, it is certainly bounded, and so the definition makes sense. Moreover, the norm properties follow, essentially from those of $\|\cdot\|_\infty$ on $\mathbb{F}^n$. This norm we call, of course, the ***∞-norm***.

Now we consider yet another generalisation of the three types of norms we have been considering, now thinking about, not sequences, but functions. The reader should note the very strong analogies between the definitions of the norms that follow and the norms above: the sums are replaced with integrals and the "max" is replaced with a "sup." Since the issues surrounding norms on infinite-dimensional vector spaces can be complex, one should cling to familiarity where possible.

8. We consider the $\mathbb{F}$-vector space $\mathsf{C}^0([a,b];\mathbb{F})$ of continuous $\mathbb{F}$-valued functions on the compact interval $[a,b]$. Provided that $b > a$ this is an infinite-dimensional vector space, cf. Example 4.3.18–**??**. On this vector space we define

$$\|f\|_2 = \Big(\int_a^b |f(x)|^2\,\mathrm{d}x\Big)^{1/2}.$$

Note that continuous functions (and therefore their squares) on compact intervals are always Riemann integrable by Corollary 3.4.12, and so the integral here is the friendly Riemann integral. It is easy to see that this possible norm satisfies the homogeneity and positive-definiteness properties of a norm (see Exercise 3.4.1 for positive-definiteness). Thus, like its 2-norm brother on $\mathbb{F}^n$, the difficult norm property to verify is the triangle inequality. However, we shall see in ***missing stuff*** that this norm is derived from an inner product, and so this will give the triangle inequality just like the 2-norm on $\mathbb{F}^n$. We shall also see this norm arise from the more general setting of Section 6.7.8. Again, the point is that we will subsequently prove the triangle inequality for $\|\cdot\|_2$ in a few different ways.

This norm will be called the ***2-norm*** on $\mathsf{C}^0([a,b];\mathbb{F})$.

9. On $\mathsf{C}^0([a,b];\mathbb{F})$ define

$$\|f\|_1 = \int_a^b |f(x)|\,\mathrm{d}x.$$

Again, the integral here is the Riemann integral. The three norm properties are easily verified. Only the triangle inequality is possibly nontrivial:

$$\begin{aligned}\|f+g\|_1 &= \int_a^b |f(x)+g(x)|\,\mathrm{d}x \le \int_a^b \big(|f(x)|+|g(x)|\big)\,\mathrm{d}x\\ &= \int_a^b |f(x)|\,\mathrm{d}x + \int_a^b |g(x)|\,\mathrm{d}x = \|f\|_1 + \|g\|_1.\end{aligned}$$

This norm, called the ***1-norm***, is different than the 2-norm. As the reader can explore in Exercise 6.1.6, for the 1- and 2-norms on $\mathbb{F}^n$, there is some sort of equivalence between these. However, for the 1- and 2-norms on $\mathsf{C}^0([a,b];\mathbb{F})$ this is no longer true. This is not perfectly obvious right now, and the reader will have to wait until ***missing stuff*** to start understanding this. But this is where we start to see how things are more complicated for infinite-dimensional vector spaces.

10. As a final norm on $\mathsf{C}^0([a,b],\mathbb{F})$ we take

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in [a,b]\}.$$

Again, the triangle inequality is the troublesome property to verify. In this case the verification goes as follows:

$$\begin{aligned}\|f+g\|_\infty &= \sup\{|f(x)+g(x)| \mid x\in[0,1]\}\\ &\le \sup\{|f(x)|+|g(x)| \mid x\in[0,1]\}\\ &\le \sup\{|f(x)|+|g(y)| \mid x,y\in[0,1]\}\\ &\le \sup\{|f(x)| \mid x\in[0,1]\}+\sup\{|g(y)| \mid y\in[0,1]\}\\ &= \|f\|_\infty+\|g\|_\infty.\end{aligned}$$

This norm is yet again different than the 1- and 2-norms. Moreover, it is yet again fundamentally not equivalent, distinguishing the infinite-dimensional case from the finite-dimensional case. This will be elucidated in ***missing stuff***. •

An obvious question is whether a vector space always possesses a norm. The answer is, "Yes, it does," and the astute reader will have seen from Examples 5, 6, and 7 above how this can be done. We record this as the following result.

6.1.4 Proposition (Vector spaces always have at least one norm) *If $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and if V is an $\mathbb{F}$-vector space then there is a norm on V.*

Proof By Theorem 4.3.46 we know the vector space V possesses a basis which establishes an isomorphism ι of V with $\mathbb{F}_0^J$ for some set J. Let us first define a norm on $\mathbb{F}_0^J$. Writing a typical element of $\mathbb{F}_0^J$ as $(v_j)_{j\in J}$ we define

$$\|(v_j)_{j\in J}\|_J=\sum_{j\in J}|v_j|,$$

noting that this sum exists since all but finitely many of the v_j's are zero. To verify that this is a norm is straightforward, cf. Example 6.1.3–5. Now define $\|\cdot\|_{\mathsf{V}}$ by $\|v\|_{\mathsf{V}}=\|\iota(v)\|_J$. That this is indeed defines a norm follows from linearity of ι:

$$\begin{aligned}&\|av\|_{\mathsf{V}}=\|\iota(av)\|_J=\|a\iota(v)\|_J=|a|\|\iota(v)\|=|a|\|v\|_{\mathsf{V}};\\ &\|v_1+v_2\|_{\mathsf{V}}=\|\iota(v_1)+\iota(v_2)\|_J\le\|\iota(v_1)\|_J+\|\iota(v_2)\|_J=\|v_1\|_{\mathsf{V}}+\|v_2\|_{\mathsf{V}}.\end{aligned}$$

Also, if $\|v\|_{\mathsf{V}}=0$ this $\|\iota(v)\|_J=0$ which means that $\iota(v)=0_{\mathbb{F}_0^J}$. Thus $v=0_{\mathsf{V}}$ since ι is an isomorphism. ■

One needs to take care with the preceding result: (1) it does not say that there is a *unique* norm on a given vector space; (2) it does not say that there is a useful norm on a given vector space. Indeed, we will see in Corollary 6.6.27 that some vector spaces do not possess "useful" norms. Thus the result should be thought of as being in the interesting vein rather than the useful vein, particularly for infinite-dimensional normed vector spaces.

In terms of convenient lingo the following definition is helpful.

6.1.5 Definition (Seminormed vector space, normed vector space) Let $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$.

(i) A ***seminormed* $\mathbb{F}$*-vector space*** is a pair $(\mathsf{V},\|\cdot\|)$ where V is a $\mathbb{F}$-vector space and $\|\cdot\|$ is a seminorm on V.

(ii) A ***normed* $\mathbb{F}$*-vector space*** is a pair $(\mathsf{V},\|\cdot\|)$ where V is a $\mathbb{F}$-vector space and $\|\cdot\|$ is a norm on V. •

6.1.6 Notation ((Semi)normed vector spaces) If a norm or seminorm is understood, we shall often say, "the (semi)normed $\mathbb{F}$-vector space V." One really needs to exercise caution with this abuse, however, since the same vector space can have multiple norms, and the behaviour can depend in a drastic way on the norm. •

Let us give some more or less trivial properties of normed vector spaces.

6.1.7 Proposition (Properties of seminormed and normed vector spaces) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|)$ be a seminormed $\mathbb{F}$-vector space, and let $\mathsf{U} \subseteq \mathsf{V}$ be a subspace. Then the following statements hold:*

(i) the map $(\mathrm{v}_1, \mathrm{v}_2) \mapsto \|\mathrm{v}_1 - \mathrm{v}_2\|$ is a semimetric on V, and is a metric when $\|\cdot\|$ is a norm;

(ii) $\big|\|\mathrm{v}_1\| - \|\mathrm{v}_2\|\big| \le \|\mathrm{v}_1 - \mathrm{v}_2\|$ for all $\mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V}$;

(iii) $\big|\|\mathrm{v}_1 - \mathrm{v}_3\| - \|\mathrm{v}_2 - \mathrm{v}_4\|\big| \le \|\mathrm{v}_1 - \mathrm{v}_2\| + \|\mathrm{v}_3 - \mathrm{v}_4\|$ for all $\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3, \mathrm{v}_4 \in \mathsf{V}$;

(iv) the restriction of $\|\cdot\|$ to U defines a seminorm on U, and this seminorm is a norm when $\|\cdot\|$ is a norm.

Proof (i) This is just a matter of plugging in the definitions. Perhaps the only nontrivial fact is the triangle inequality:

$$\|v_1 - v_3\| = \|(v_1 - v_2) + (v_2 - v_3)\| \le \|v_1 - v_2\| + \|v_2 - v_3\|.$$

(ii) This is Exercise 6.1.3.

(iii) We use the triangle inequality and part (ii):

$$\begin{aligned}\big|\|v_1 - v_3\| - \|v_2 - v_4\|\big| &= \big|\|v_1 - v_3\| - \|v_2 - v_3\| + \|v_3 - v_2\| - \|v_2 - v_4\|\big| \\ &\le \big|\|v_1 - v_3\| - \|v_2 - v_3\|\big| + \big|\|v_3 - v_2\| + \|v_2 - v_4\|\big| \\ &\le \|v_1 - v_2\| + \|v_3 - v_4\|,\end{aligned}$$

as desired.

(iv) This is trivial. ■

Now we indicate how one can pass from a seminormed vector space to a normed vector space in a natural way. This mirrors our result Theorem **??** for semimetric spaces.

6.1.8 Theorem (Normed vector spaces from seminormed vector spaces) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a seminormed $\mathbb{F}$-vector space. Then the following statements hold:*

(i) the set $\mathsf{V}_0 = \{\mathrm{v} \in \mathsf{V} \mid \|\mathrm{v}\| = 0\}$ is a subspace of V;

(ii) the function $\mathsf{V}/\mathsf{V}_0 \ni \mathrm{v} + \mathsf{V}_0 \mapsto \|\mathrm{v}\|$ is a norm on V/V_0.

Proof (i) If $u, v \in \mathsf{V}_0$ and if $a \in \mathbb{F}$ then

$$0 \le \|u + v\| \le \|u\| + \|v\| = 0$$

and

$$\|av\| = |a|\|v\| = 0,$$

giving $u + v, av \in \mathsf{V}_0$, as desired.

(ii) First let us show that the function is well-defined. Suppose that $v + \mathsf{V}_0 = v' + \mathsf{V}_0$ so that $v - v' \in \mathsf{V}_0$. Then

$$\|v'\| = \|v + (v' - v)\| \leq \|v\| + \|v' - v\| = \|v\|$$

and

$$\|v\| = \|v' + (v - v')\| \leq \|v'\| + \|v - v'\| = \|v'\|$$

using the triangle inequality. Thus $\|v'\| = \|v\|$, and the map is then well-defined. It clearly has the homogeneity and triangle inequality properties of a norm. To check the positive-definiteness, suppose that $\|v + \mathsf{V}_0\| = 0$. Then $\|v\| = 0$ and so $v \in \mathsf{V}_0$, giving $v + \mathsf{V}_0 = 0_{\mathsf{V}} + \mathsf{V}_0$, as desired. ■

6.1.2 Open and closed subsets of normed vector spaces

As we saw in Proposition 6.1.7 a seminorm (resp. norm) $\|\cdot\|$ on V determines a semimetric (resp. metric) on V by $d_{\|\cdot\|}(v_1, v_2) = \|v_1 - v_2\|$. A semimetric then determines a topology, and, if the semimetric is a metric, this topology is Hausdorff (see ***missing stuff***). Therefore, seminormed vector spaces are topological spaces, and normed vector spaces are Hausdorff topological spaces. In this section we describe this topology in more detail. Some of what we say is redundant since it follows from what we have already said for metric spaces. However, we aim to make our treatment of normed vector spaces as self-contained as possible. In this section we make statements that are valid for seminormed vector spaces, and not just normed vector spaces, although it is the latter that are of most immediate interest. We adopt the convention of writing "(semi)norm" when we mean that the object can be either a norm or a seminorm. Readers caring only about norms can omit the "(semi)" in their heads.

As with metrics, the building block of the topology of a normed vector space is the open ball.

6.1.9 Definition (Open, closed, and bounded sets in (semi)normed vector spaces) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space.

(i) The ***open ball*** of radius r about $v_0 \in \mathsf{V}$ is the set

$$\mathsf{B}(r, v_0) = \{v \in \mathsf{V} \mid \|v - v_0\| < r\}.$$

(ii) The ***closed ball*** of radius r about $v_0 \in \mathsf{V}$ is the set

$$\overline{\mathsf{B}}(r, v_0) = \{v \in \mathsf{V} \mid \|v - v_0\| \leq r\}.$$

(iii) A subset $U \subseteq \mathsf{V}$ is ***open*** if, for each $v \in U$, there exists $\epsilon \in \mathbb{R}_{>0}$ such that $\mathsf{B}(\epsilon, v) \subseteq U$. (The empty set is also open, by declaration.)

(iv) A subset $A \subseteq \mathsf{V}$ is ***closed*** is $\mathsf{V} \setminus A$ if open.

(v) A subset $A \subseteq \mathsf{V}$ is ***bounded*** if there exists $R \in \mathbb{R}_{>0}$ such that $A \subseteq \mathsf{B}(R, 0_{\mathsf{V}})$. •

One can easily show that the open ball is open (this is Exercise 6.1.1).

We shall not attempt to systematically distinguish notationally the rôle of $\|\cdot\|$ in the open ball $\mathsf{B}(r, v_0)$. If there is a potential cause of confusion we will handle it as it comes up. For example, if we are working with multiple (semi)normed vector spaces, we may use the notation $\mathsf{B}_{\mathsf{V}}(r, v_0)$ to specify that a ball is in V.

Let us give some properties of open sets.

6.1.10 Proposition (Properties of open subsets of (semi)normed vector spaces) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a (semi)normed* $\mathbb{F}$*-vector space. Then the following statements hold:*

(i) for $(\mathrm{U}_a)_{a\in A}$ *an arbitrary family of open sets,* $\cup_{a\in A}\mathrm{U}_a$ *is open;*

(ii) for $(\mathrm{U}_1, \dots, \mathrm{U}_n)$ *a finite family of open sets,* $\cap_{j=1}^{n}\mathrm{U}_j$ *is open.*

Proof (i) Let $v \in \cup_{a\in A} U_a$. Then, since $v \in U_{a_0}$ for some $a_0 \in A$, there exists $\epsilon \in \mathbb{R}_{>0}$ such that $\mathsf{B}(\epsilon, v) \subseteq U_{a_0} \subseteq \cup_{a\in A} U_a$.

(ii) Let $v \in \cap_{j=1}^{n} U_j$. For each $j \in \{1, \dots, n\}$, choose $\epsilon_j \in \mathbb{R}_{>0}$ such that $\mathsf{B}(\epsilon_j, v) \subseteq U_j$, and let $\epsilon = \min\{\epsilon_1, \dots, \epsilon_n\}$. Then $\mathsf{B}(\epsilon, v) \subseteq U_j$, $j \in \{1, \dots, n\}$, and so $\mathsf{B}(\epsilon, v) \subseteq \cap_{j=1}^{n} U_j$. ■

This result shows that the collection of open subsets of a (semi)normed vector space define a topology.

6.1.11 Definition ((Semi)norm topology) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. The topology on V whose open sets are the open sets defined by the (semi)norm $\|\cdot\|$ is the ***(semi)norm topology*** on V. •

One of the most important properties about the norm topology is that it is translation invariant. Let us see what this means. For $v_0 \in \mathsf{V}$ define $\tau_{v_0} \colon \mathsf{V} \to \mathsf{V}$ by $\tau_{v_0}(v) = v + v_0$. Thus τ_{v_0} is "translation by v_0." We then have the following result.

6.1.12 Proposition (Translation invariance of the (semi)norm topology) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a (semi)normed* $\mathbb{F}$*-vector space. Then a subset* $\mathrm{U} \subseteq \mathsf{V}$ *is open if and only if* $\tau_{v_0}(\mathrm{U})$ *is open.*

Proof Suppose that $U \subseteq \mathsf{V}$ is open and let $v \in \tau_{v_0}(U)$. Then $\tau_{-v_0}(v) \in U$ and so there exists $\epsilon > 0$ such that $\mathsf{B}(\epsilon, v) \subseteq U$. Note that

$$\begin{aligned}\tau_{v_0}(\mathsf{B}(\epsilon, v)) &= \tau_{v_0}(\{u \in \mathsf{V} \mid \|u - v\| < \epsilon\}) \\ &= \{v_0 + u \in \mathsf{V} \mid \|u - v\| < \epsilon\} \\ &= \{u' \in \mathsf{V} \mid \|u' - (v + v_0)\| < \epsilon\} \\ &= \mathsf{B}(\epsilon, \tau_{v_0}(v)).\end{aligned}$$

Thus

$$\mathsf{B}(\epsilon, v) \subseteq U \implies \tau_{v_0}(\mathsf{B}(\epsilon, v)) \subseteq \tau_{v_0}(U) \implies \mathsf{B}(\epsilon, \tau_{v_0}(v)) \subseteq \tau_{v_0}(U).$$

Thus $\tau_{v_0}(U)$ is open.

Conversely, if $\tau_{v_0}(U)$ is open then, by the first part of the proof, $\tau_{-v_0}(\tau_{v_0}(U)) = U$ is open. ■

As the proof of the preceding result makes clear, the key to the translation invariance of the norm topology is the fact that $\tau_{v_0}(\mathsf{B}(r, v)) = \mathsf{B}(r, \tau_{v_0}(v))$ for every

$r \in \mathbb{R}_{>0}$ and $v, v_0 \in \mathsf{V}$. This is a pretty obvious fact, but is so useful that it is worth pointing out explicitly.

The norm topology generally depends on the norm. However, it is possible that two different norms will give the same topology. The following definition captures this idea.

6.1.13 Definition (Equivalent norms) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be a $\mathbb{F}$-vector space. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ (the subscripts "1" and "2" have nothing to do with the 1- and 2-norms considered in Example 6.1.3) are ***equivalent*** if a subset $U \subseteq \mathsf{V}$ is open in the norm topology defined by $\|\cdot\|_1$ if and only if it is open in the norm topology defined by $\|\cdot\|_2$. •

We will not be interested in the notion of equivalence for seminorms.

In short, equivalent norms define the same open sets. It is useful to be able to characterise equivalent norms in a more computational manner, one that might be able to check in practice. The following result gives just such a characterisation.

6.1.14 Theorem (Characterisation of equivalent norms) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be a $\mathbb{F}$-vector space. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on V are equivalent if and only if there exists $\mathrm{C} \in \mathbb{R}_{>0}$ such that*

$$\mathrm{C}^{-1}\|\mathrm{v}\|_2 \leq \|\mathrm{v}\|_1 \leq \mathrm{C}\|\mathrm{v}\|_2$$

for all $\mathrm{v} \in \mathsf{V}$.

Proof First suppose that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. Let $\mathsf{B}_1(r, v_0)$ and $\mathsf{B}_2(r, v_0)$ denote the open balls of radius r centred at v_0 for $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively. By ***missing stuff***, equivalence of the two norm topologies implies that for every $R \in \mathbb{R}_{>0}$ there exists $C_1, C_2 \in \mathbb{R}_{>0}$ such that

$$\mathsf{B}_2(C_1, 0_{\mathsf{V}}) \subseteq \mathsf{B}_1(R, 0_{\mathsf{V}}) \subseteq \mathsf{B}_2(C_2, 0_{\mathsf{V}}).$$

Let us consider the inclusion $\mathsf{B}_2(C_1, 0_{\mathsf{V}}) \subseteq \mathsf{B}_1(R, 0_{\mathsf{V}})$. If $v \in \mathsf{V}$ is nonzero then this inclusion gives

$$\begin{aligned} \|v\|_2 \leq 1 \implies \|C_1 v\|_2 \leq C_1 \implies \|C_1 v\|_1 \leq R \implies \frac{\|C_1 v\|_1}{\|C_1 v\|_2} \leq \frac{R}{\|C_1 v\|_2} \\ \implies \frac{\|v\|_1}{\|v\|_2} \leq \frac{R}{C_1} \implies \|v\|_1 \leq \frac{R}{C_1}\|v\|_2. \end{aligned}$$

Thus $\|v\|_1 \leq \frac{R}{C_1}\|v\|_2$ holds if v is nonzero and if $\|v\|_2 \leq 1$. Clearly the same equality holds for $v = 0_{\mathsf{V}}$. For $v \in \mathsf{V}$ nonzero we also have

$$\Big\|\frac{v}{\|v\|_2}\Big\|_1 \leq \frac{R}{C_1}\Big\|\frac{v}{\|v\|_2}\Big\|_2 \implies \|v\|_1 \leq \frac{R}{C_1}\|v\|_2.$$

Thus the relation $\|v\|_1 \leq \frac{R}{C_1}\|v\|_2$ holds for all $v \in \mathsf{V}$.

An entirely similar argument shows that the inclusion $\mathsf{B}_1(R, 0_{\mathsf{V}}) \subseteq \mathsf{B}_2(C_2, 0_{\mathsf{V}})$ implies that $\|v\|_2 \leq \frac{C_2}{R}\|v\|_1$ for all $v \in \mathsf{V}$. Thus we have

$$\frac{C_2}{R}\|v\|_2 \leq \|v\|_1 \leq \frac{R}{C_1}\|v\|_2$$

for all $v \in \mathsf{V}$. Taking $C = \max\{\frac{R}{C_1}, \frac{R}{C_2}\}$ gives

$$C^{-1}\|v\|_2 \leq \|v\|_1 \leq C\|v\|_2, \qquad v \in \mathsf{V},$$

as desired.

Now suppose that there exists $C \in \mathbb{R}_{>0}$ such that

$$C^{-1}\|v\|_2 \leq \|v\|_1 \leq C\|v\|_2$$

for all $v \in \mathsf{V}$. Let $R \in \mathbb{R}_{>0}$ and note that

$$v \in \mathsf{B}_1(R, 0_{\mathsf{V}}) \implies \|v\|_1 < R \implies \|v\|_2 \leq RC \implies v\mathsf{B}_2(RC, 0_{\mathsf{V}}).$$

Thus $\mathsf{B}_1(R, 0_{\mathsf{V}}) \subseteq \mathsf{B}_2(RC, 0_{\mathsf{V}})$. Similarly we show that $\mathsf{B}_2(\frac{R}{C}, 0_{\mathsf{V}}) \subseteq \mathsf{B}_1(R, 0_{\mathsf{V}})$. Thus we have

$$\mathsf{B}_2(\tfrac{R}{C}, 0_{\mathsf{V}}) \subseteq \mathsf{B}_1(R, 0_{\mathsf{V}}) \subseteq \mathsf{B}_2(RC, 0_{\mathsf{V}})$$

for every $R \in \mathbb{R}_{>0}$. From the remarks following the proof of Proposition 6.1.12 it follows that

$$\mathsf{B}_2(\tfrac{R}{C}, v_0) \subseteq \mathsf{B}_1(R, v_0) \subseteq \mathsf{B}_2(RC, v_0)$$

for every $R \in \mathbb{R}_{>0}$ and every $v_0 \in \mathsf{V}$. The equivalence of the two norm topologies now follows from ***missing stuff***. ■

The following result shows that, on a finite-dimensional normed vector space there is really only one norm topology, although one can use different norms to define it.

6.1.15 Theorem (Uniqueness of the norm topology on finite-dimensional normed vector spaces) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if V is a finite-dimensional $\mathbb{F}$-vector space, then any two norms on V are equivalent.*

Proof Let $\{e_1, \dots, e_n\}$ be a basis for V and let $\iota\colon \mathsf{V} \to \mathbb{F}^n$ be defined by

$$\iota(v_1e_1, \dots, v_ne_n) = (v_1, \dots, v_n).$$

Define norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on V by

$$\|v\|_1 = \|\iota(v)\|_1 = \sum_{j=1}^{n} |v_j|,$$
$$\|v\|_2 = \|\iota(v)\|_2 = \Big(\sum_{j=1}^{n} |v_j|^2\Big)^{1/2}.$$

Thus we are abusing notation and using $\|\cdot\|_1$ and $\|\cdot\|_2$ for norms both on V and $\mathbb{F}^n$. These do define norms on V by Example 6.1.3 (also, cf. the proof of Proposition 6.1.4). Since the notion of equivalence of norms is an equivalence relation (this is Exercise 6.1.5), it suffices to show that any other norm of V is equivalent to $\|\cdot\|_2$. Let $\|\cdot\|$ be another norm on V and write, for $u, v \in \mathsf{V}$,

$$u = u_1e_1 + \dots + u_ne_n, \quad v = v_1e_1 + \dots + v_ne_n.$$

We then have, by Exercise 6.1.3 and Proposition **??**, and the triangle inequality,

$$\begin{aligned}|\|u\| - \|v\|| \leq \|u - v\| = \Big\|\sum_{j=1}^{n}(u_j - v_j)e_j\Big\| &\leq \sum_{j=1}^{n}|u_j - v_j|\|e_j\| \\ &\leq \max\{\|e_j\| \mid j \in \{1, \dots, n\}\}\|u - v\|_1 \leq C\|v\|_2,\end{aligned}$$

where $C = \max\{\|e_j\| \mid j \in \{1, \dots, n\}\}\sqrt{n}$. We claim that this implies that the function $v \mapsto \|\iota^{-1}(v)\|$ on $\mathbb{F}^n$ is continuous with respect to the norm $\|\cdot\|_2$. Indeed, for $\epsilon \in \mathbb{R}_{>0}$ let $\delta = \frac{\epsilon}{C}$. For $v_0 \in \mathbb{F}^n$ suppose that $\|v - v_0\|_2 < \delta$. Then, from our computations above,

$$|\|\iota^{-1}(v)\| - \|\iota^{-1}(v_0)\|| \leq C\|v - v_0\|_2 < \epsilon,$$

giving continuity of $v \mapsto \|\iota^{-1}(v)\|$ at v_0. Let $\overline{\mathsf{B}}_2(1, \mathbf{0}_{\mathbb{F}^n})$ be the unit ball with respect to the norm $\|\cdot\|_2$ centred at the origin in $\mathbb{F}^n$ and let $\overline{\mathsf{B}}_2(1, 0_{\mathsf{V}})$ be the unit ball with respect to the norm $\|\cdot\|_2$ centred at the origin in V. The boundary of $\overline{\mathsf{B}}_2(1, \mathbf{0}_{\mathbb{F}^n})$ is closed and bounded with respect to the norm $\|\cdot\|_2$ and its topology, and so is compact in $\mathbb{F}^n$ with respect to the usual topology by the Heine–Borel Theorem. Therefore, by ***missing stuff***, the function $v \mapsto \|\iota^{-1}(v)\|$ attains a minimum value $m \in \mathbb{R}_{>0}$ and a maximum value $M \in \mathbb{R}_{>0}$ on $\mathrm{bd}(\overline{\mathsf{B}}_2(1, \mathbf{0}_{\mathbb{F}^n}))$. Thus, for $v \in \overline{\mathsf{B}}_2(1, \mathbf{0}_{\mathbb{F}^n})$ we have

$$m \leq \|\iota^{-1}(v)\| \leq M$$

which is equivalent to saying that, for $v \in \mathrm{bd}(\overline{\mathsf{B}}_2(1, 0_{\mathsf{V}}))$ (boundary being taken with respect to the norm topology on V for the norm $\|\cdot\|_2$) we have

$$m \leq \|v\| \leq M.$$

For arbitrary $v \in \mathsf{V} \setminus \{0_{\mathsf{V}}\}$ this gives

$$m \leq \Big\|\frac{v}{\|v\|_2}\Big\| \leq M \quad \implies \quad m\|v\|_2 \leq \|v\| \leq M\|v\|_2,$$

showing that $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent if we take $C = \max\{M, m^{-1}\}$. ■

We will use this theorem to unambiguously talk about the norm topology on $\mathbb{F}^n$, or any finite-dimensional $\mathbb{F}$-vector space, as being the topology defined by *any* norm.

6.1.3 Subspaces, direct sums, and quotients

We have studied in Section 4.3 the notions of subspace, direct sum, and quotient from an algebraic point of view. Let us see now how these notions interact with the structure of a norm.

For subspaces we record the following trivial result. We will have much more to say about subspaces of normed vector spaces in Section 6.6.4.

6.1.16 Proposition (Subspaces of (semi)normed vector spaces are (semi)normed vector spaces) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a (semi)normed* $\mathbb{F}$*-vector space. If* $\mathsf{U} \subseteq \mathsf{V}$ *is a subspace then the map* $\mathsf{U} \ni \mathrm{u} \mapsto \|\mathrm{u}\| \in \mathbb{R}_{\geq 0}$ *is a (semi)norm on* U*.*

Now we consider direct sums of normed vector spaces. Let us first consider the general case, and then consider the case of finite direct sums as a special case. We recall from Definition 4.3.40 that the direct sum of a family $(\mathsf{V}_i)_{i \in I}$ of vector spaces is the set of maps $\phi \colon I \to \cup_{i \in I} \mathsf{V}_i$ for which $\phi(i) \in \mathsf{V}_i$, $i \in I$, and for which the set $\{i \in I \mid \phi(i) \neq 0_{\mathsf{V}_i}\}$ is finite. This set has a natural vector space structure and is denoted $\bigoplus_{i \in I} \mathsf{V}_i$.

6.1.17 Theorem (Direct sums of (semi)normed vector spaces are (semi)normed vector spaces) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $((\mathsf{V}_\mathrm{i}, \|\cdot\|_\mathrm{i}))_{\mathrm{i} \in \mathrm{I}}$ *be a family of (semi)normed* $\mathbb{F}$*-vector spaces. For* $\phi \in \bigoplus_{\mathrm{i} \in \mathrm{I}} \mathsf{V}_\mathrm{i}$ *define*

$$\|\phi\|_\mathrm{I} = \sum_{\mathrm{i} \in \mathrm{I}} \|\phi(\mathrm{i})\|_\mathrm{i},$$

this sum being well-defined since it is finite. Then $(\bigoplus_{\mathrm{i} \in \mathrm{I}} \mathsf{V}_\mathrm{i}, \|\cdot\|_\mathrm{I})$ *is a (semi)normed* $\mathbb{F}$*-vector space, and is moreover a normed vector space if each of the components* $(\mathsf{V}_\mathrm{i}, \|\cdot\|_\mathrm{i})$*,* $\mathrm{i} \in \mathrm{I}$*, is a normed vector space.*

Proof Let $a \in \mathbb{F}$ and compute

$$\|a\phi\|_I = \sum_{i \in I} \|a\phi(i)\|_a = \sum_{i \in I} |a| \|\phi(i)\|_i = |a| \sum_{i \in I} \|\phi(i)\|_a = |a| \|\phi\|_I,$$

where all operations make sense since the sums are finite.

If $\phi, \psi \in \bigoplus_{i \in I} \mathsf{V}_i$ we compute

$$\|\phi + \psi\|_I = \sum_{i \in I} \|\phi(i) + \psi(i)\|_i \leq \sum_{i \in I} \|\phi(i)\|_i + \sum_{i \in I} \|\psi(i)\|_i = \|\phi\|_I + \|\psi\|_I,$$

as desired.

Finally, if

$$\|\phi\| = \sum_{i \in I} \|\phi(i)\|_i = 0$$

then we must have $\|\phi(i)\|_i = 0$ for each $i \in I$. If each of the seminorms $\|\cdot\|_i$, $i \in I$, are norms then this implies that $\phi(i) = 0_{\mathsf{V}_i}$, $i \in I$, implying that $\|\cdot\|_I$ is a norm. ■

We can now make the following definition.

6.1.18 Definition (Direct sum of (semi)normed vector spaces) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $((\mathsf{V}_i, \|\cdot\|_i))_{i \in I}$ be a family of (semi)normed $\mathbb{F}$-vector spaces. The (semi)normed vector space $(\bigoplus_{i \in I} \mathsf{V}_i, \|\cdot\|_I)$ is the ***direct sum*** of $((\mathsf{V}_i, \|\cdot\|_i))_{i \in I}$. •

Let us record how this works for the direct sum of two (semi)normed vector spaces. Thus let $(\mathsf{V}_1, \|\cdot\|_1)$ and $(\mathsf{V}_2, \|\cdot\|_2)$ be (semi)normed $\mathbb{F}$-vector spaces. Their direct sum is the vector space $\mathsf{V}_1 \oplus \mathsf{V}_2$, points in which we denote by (v_1, v_2), with the (semi)norm

$$\|(v_1, v_2)\|_{1,2} = \|v_1\|_1 + \|v_2\|_2.$$

Now we consider quotients of normed vector spaces by subspaces.

6.1.19 Proposition (The quotient of a (semi)normed vector space is a (semi)normed vector space) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space, and let U be a subspace. If we define*

$$\|\mathrm{v} + \mathsf{U}\|_{/\mathsf{U}} = \inf\{\|\mathrm{v} + \mathrm{u}\| \mid \mathrm{u} \in \mathsf{U}\}$$

then $\|\cdot\|_{/\mathsf{U}}$ is a seminorm on V/U. Moreover, if $\|\cdot\|$ is a norm and if U is closed, then $\|\cdot\|_{/\mathsf{U}}$ is a norm.

Proof It is evident that $\|v + \mathsf{U}\|_{/\mathsf{U}} \in \mathbb{R}_{>0}$. If $a = 0$ we have

$$\|0(v + \mathsf{U})\|_{/\mathsf{U}} = \|0v + \mathsf{U}\|_{/\mathsf{U}} = \inf\{\|0_{\mathsf{V}} + u\| \mid u \in \mathsf{U}\} = 0 = |a|\|v + \mathsf{U}\|_{/\mathsf{U}}.$$

For $a \in \mathbb{F} \setminus \{0\}$ we have

$$\begin{aligned}\|a(v + \mathsf{U})\|_{/\mathsf{U}} = \|av + \mathsf{U}\|_{/\mathsf{U}} &= \inf\{\|av + u\| \mid u \in \mathsf{U}\}\\ &= \inf\{\|av + au'\| \mid u' \in \mathsf{U}\} = \inf\{|a|\|v + u'\| \mid u' \in \mathsf{U}\}\\ &= |a| \inf\{\|v + u'\| \mid u' \in \mathsf{U}\} = |a|\|v + \mathsf{U}\|_{/\mathsf{U}}.\end{aligned}$$

For the triangle inequality we have

$$\begin{aligned}\|(v_1 + \mathsf{U}) + (v_2 + \mathsf{U})\|_{/\mathsf{U}} &= \|(v_1 + v_2) + \mathsf{U}\|_{/\mathsf{U}} = \inf\{\|v_1 + v_2 + u\| \mid u \in \mathsf{U}\}\\ &= \inf\{\|v_1 + v_2 + u_1 + u_2\| \mid u_1, u_2 \in \mathsf{U}\}\\ &\le \inf\{\|v_1 + u_1\| + \|v_2 + u_2\| \mid u_1, u_2 \in \mathsf{U}\}\\ &= \inf\{\|v_1 + u_1\| \mid u_1 \in \mathsf{U}\} + \inf\{\|v_2 + u_2\| \mid u_2 \in \mathsf{U}\}\\ &= \|v_1 + \mathsf{U}\|_{/\mathsf{U}} + \|v_2 + \mathsf{U}\|_{/\mathsf{U}},\end{aligned}$$

as desired, where we have used Proposition 2.2.27.

To prove the final assertion we rely on some facts about closed sets that we will not prove until Section 6.6.2. Let $v + \mathsf{U} \in \mathsf{V}/\mathsf{U}$ satisfy $\|v + \mathsf{U}\|_{/\mathsf{U}} = 0$. Thus

$$\inf\{\|v + u\| \mid u \in \mathsf{U}\} = 0.$$

Therefore, for $j \in \mathbb{Z}_{>0}$, there exists $u_j \in \mathsf{U}$ such that $\|v + u_j\| < \frac{1}{j}$. Thus the sequence $(v + u_j)_{j \in \mathbb{Z}_{>0}}$ converges to 0_{V}. By Proposition 6.2.6 it follows that the sequence $(u_j)_{j \in \mathbb{Z}_{>0}}$ converges to $-v$. Since the sequence is in U and since U is closed, by Proposition 6.6.8(ii) it follows that $-v \in \mathsf{U}$ and so $v \in \mathsf{U}$. Thus $v + \mathsf{U} = 0_{\mathsf{V}} + \mathsf{U}$, giving $\|\cdot\|_{/\mathsf{U}}$ as a norm. ■

One should be a little careful with the result. It does not say that $\|\cdot\|_{/\mathsf{U}}$ is a norm if $\|\cdot\|$ is a norm; this requires the additional assumption that U is closed.

Let us examine some properties of the canonical projection from V to V/U.

Let us examine some properties of the canonical projection from V to V/U. Here we refer ahead to Section 6.5 for notion of continuity and back to ***missing stuff*** for the notion of the quotient topology.

6.1.20 Proposition (The canonical projection onto the quotient is continuous) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space, and let U be a subspace. Then the canonical projection $\pi_{\mathsf{U}} \colon \mathsf{V} \to \mathsf{V}/\mathsf{U}$ is continuous. Moreover, the seminorm topology on V/U coincides with the quotient topology.*

Proof Let $v \in \mathsf{V}$ with $v + \mathsf{U}$ the projection to V/U. Let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in V converging to v. We claim that $(v_j + \mathsf{U})_{j\in\mathbb{Z}_{>0}}$ converges to $v + \mathsf{U}$. Indeed, if $\epsilon \in \mathbb{R}_{>0}$, take $N \in \mathbb{Z}_{>0}$ such that $\|v - v_j\| < \epsilon$ for $j \geq N$. Then

$$\|(v - v_j) + \mathsf{U}\|_{/\mathsf{U}} \leq \|v - v_j\| < \epsilon$$

for $j \geq N$, giving convergence as desired. Continuity of $v \mapsto v + \mathsf{U}$ now follows from Theorem 6.5.2.

Let $\pi\colon \mathsf{V} \to \mathsf{V}/\mathsf{U}$ denote the canonical projection. Now let $S \subseteq \mathsf{V}/\mathsf{U}$ be such that $\pi^{-1}(S)$ is a open. We claim that S is a open. For $v_0 + \mathsf{U} \in S$ let $\mathsf{B}_{\mathsf{V}}(\epsilon, v_0)$ be an open ball about v_0 contained in $\pi^{-1}(S)$. We have

$$\begin{aligned}\pi(\mathsf{B}_{\mathsf{V}}(\epsilon, v_0)) &= \{v + \mathsf{U} \mid \|v - v_0\| < \epsilon\} = \{v + \mathsf{U} \mid \|(v - v_0) + \mathsf{U}\|_{/\mathsf{U}} < \epsilon\} \\ &= \mathsf{B}_{\mathsf{V}/\mathsf{U}}(\epsilon, v_0 + \mathsf{U}).\end{aligned}$$

Since $\pi(\mathsf{B}_{\mathsf{V}}(\epsilon, v_0)) \subseteq S$ it follows that S is open. ■

Exercises

6.1.1 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Show that $\mathsf{B}(r, v_0)$ is open for every $r \in \mathbb{R}_{>0}$ and $v_0 \in \mathsf{V}$.

6.1.2 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a $\mathbb{F}$-vector space. Let $r_1, r_2 \in \mathbb{R}_{>0}$ satisfy $r_2 \leq r_1$ and let $v_1, v_2 \in \mathbb{R}^n$. Show that if $\overline{\mathsf{B}}(r_1, v_1) \cap \overline{\mathsf{B}}(r_2, v_2) \neq \emptyset$ then $\overline{\mathsf{B}}(r_2, v_2) \subseteq \overline{\mathsf{B}}(3r_1, v_1)$. Show that you understand your proof by drawing a picture.

6.1.3 In a normed vector space $(\mathsf{V}, \|\cdot\|)$ show that for each $v_1, v_2 \in \mathsf{V}$, $|\|v_1\| - \|v_2\|| \leq \|v_1 - v_2\|$.

6.1.4 Denote by $\mathsf{C}^1([0, 1]; \mathbb{R})$ the set of $\mathbb{R}$-valued functions on $[0, 1]$ which are continuously differentiable, derivatives at 0 and 1 being taken from the right and left, respectively.

(a) For $f \in \mathsf{C}^1([0, 1]; \mathbb{R})$ define

$$\|f\| = \int_0^1 |f'(x)|\,\mathrm{d}x.$$

Show that $\|\cdot\|$ is a seminorm on $\mathsf{C}^1([0, 1]; \mathbb{R})$, but not a norm.

(b) For $f \in \mathsf{C}^1([0, 1]; \mathbb{R})$ define

$$\|f\| = |f(0)| + \int_0^1 |f'(x)|\,\mathrm{d}x.$$

Show that $\|\cdot\|$ is a norm on $\mathsf{C}^1([0, 1]; \mathbb{R})$.

6.1.5 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be an $\mathbb{F}$-vector space. Define a relation $\sim$ on the set of norms on V by saying that $\|\cdot\|_1 \sim \|\cdot\|_2$ if $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent in the sense of Definition 6.1.13. Show that $\sim$ is an equivalence relation.

6.1.6 On $\mathsf{V} = \mathbb{R}^2$ consider the three norms $\|\cdot\|_2$, $\|\cdot\|_1$, and $\|\cdot\|_\infty$ given by Examples 6.1.3–2, 6.1.3–3, and 6.1.3–4, respectively.

(a) Draw the subsets $\mathsf{B}_2(r,\mathbf{0})$, $\mathsf{B}_1(r,\mathbf{0})$, and $\mathsf{B}_\infty(r,\mathbf{0})$ of $\mathbb{R}^2$ defined by

$$\begin{aligned}\overline{\mathsf{B}}_2(r,\mathbf{0}) &= \{\boldsymbol{v}\in\mathbb{R}^2 \mid \|\boldsymbol{v}\|_2 \le r\}\\ \overline{\mathsf{B}}_1(r,\mathbf{0}) &= \{\boldsymbol{v}\in\mathbb{R}^2 \mid \|\boldsymbol{v}\|_1 \le r\}\\ \overline{\mathsf{B}}_\infty(r,\mathbf{0}) &= \{\boldsymbol{v}\in\mathbb{R}^2 \mid \|\boldsymbol{v}\|_\infty \le r\}.\end{aligned}$$

(b) Using your drawings from part (a), argue that if and only if a sequence of points $(\boldsymbol{v}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}^2$ converges in one of the three norms, it converges in the other two norms.

6.1.7 Let $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and let $(\mathsf{V},\|\cdot\|)$ be a finite-dimensional normed $\mathbb{F}$-vector space. Let $\{e_1,\dots,e_n\}$ be a basis for V for which $\|e_1\|=\cdots=\|e_n\|=1$.

(a) For $v=v_1e_1+\cdots+v_ne_n\in\mathsf{V}$ define

$$\|v\|_1=|v_1|+\cdots+|v_n|.$$

Show that $\|\cdot\|_1$ is a norm on V that satisfies $\|e_1\|_1=\cdots=\|e_n\|_1=1$.

(b) Let $\mathsf{B}(1,0_{\mathsf{V}})$ and $\mathsf{B}_1(1,0_{\mathsf{V}})$ be the unit balls for the norms $\|\cdot\|$ and $\|\cdot\|_1$, respectively. Show that $\mathsf{B}_1(1,0_{\mathsf{V}})\subseteq\mathsf{B}(1,0_{\mathsf{V}})$.

(The point is that the balls in the norm $\|\cdot\|_1$ are the smallest among the balls for all norms in which the basis vectors have unit length.)

Section 6.2

Sequences in normed vector spaces

Much of the structure of normed vector spaces can be captured by studying sequences in these spaces. Much of the presentation here follows the presentation of Section 2.3. Indeed, many of the proofs are mere changes of notation of the analogous proofs for sequences in $\mathbb{R}$. However, we give all of the details of the presentation here for both (1) completeness and (2) because not all results are *exactly* the same as those for $\mathbb{R}$. This has the disadvantage of repetitiveness, but the advantage of making this section more self-contained.

Do I need to read this section? The ideas in this section are basic, so the definitions should be read and the results understood. Readers who are familiar with the material in Section 2.3 will find this section reads pretty easily. •

6.2.1 Definitions and properties of sequences

Let V be a $\mathbb{F}$-vector space. A sequence in V is, in accordance with Definition 1.4.8, a map from $\mathbb{Z}_{>0}$ to V, and we denote a sequence by $(v_j)_{j\in\mathbb{Z}_{>0}}$. For sequences we have the usual definitions corresponding to notions of convergence.

6.2.1 Definition (Convergence of sequences) Let$\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $(\mathsf{V},\|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}$ and let $v_0 \in \mathsf{V}$. The sequence:

(i) is a ***Cauchy sequence*** if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $\|v_j - v_k\| < \epsilon$ for $j,k \geq N$;

(ii) ***converges to* $\mathbf{v_0}$** if, for each $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $\|v_j - v_0\| < \epsilon$ for $j \geq N$;

(iii) ***diverges*** if it does not converge to any element in V;

(iv) is ***bounded*** if there exists $M \in \mathbb{R}_{>0}$ such that $\|v_j\| < M$ for each $j \in \mathbb{Z}_{>0}$.

If the sequence converges to v_0 then v_0 is the ***limit*** of the sequence and we write $v_0 = \lim_{j\to\infty} v_j$. •

6.2.2 Notation (Limits with general index sets) As in Section 2.3.7 we can talk about limits of things more general than sequences. The setup where we will use this idea is the following. Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $(\mathsf{U},\|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V},\|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. We consider an open subset $O \subseteq \mathsf{U}$ and a map $\phi\colon O \to \mathsf{V}$. For $u_0 \in O$, we wish to define what we mean by $\lim_{u\to u_0}\phi(u)$. What we mean is this. If, there exists $v_0 \in \mathsf{V}$ such that, for any sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$ converging to u_0, the sequence $(\phi(u_j))_{j\in\mathbb{Z}_{>0}}$ converges to $_0$, then we write $\lim_{u\to u_0}\phi(u_0) = v_0$. •

As for sequences in $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, convergent sequences are Cauchy.

6.2.3 Proposition (Convergent sequences are Cauchy) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. If $(\mathrm{v}_j)_{j\in\mathbb{Z}_{>0}}$ is a sequence converging to v_0 then it is a Cauchy sequence.*

Proof Let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $|v_j - v_0| < \frac{\epsilon}{2}$ for $j \geq N$. Then, for $j, k \geq N$ we have

$$\|v_j - v_k\| = \|v_j - v_0 - v_k + v_0\| = \|v_j - v_0\| + \|v_k - v_0\| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,$$

using the triangle inequality. ■

Cauchy sequences are bounded.

6.2.4 Proposition (Cauchy sequences are bounded) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. If $(\mathrm{v}_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence, then it is bounded.*

Proof Choose $N \in \mathbb{Z}_{>0}$ such that $\|v_j - v_k\| < 1$ for $j, k \in \mathbb{Z}_{>0}$. Then take M_N to be the largest of the nonnegative real numbers $\|v_1\|, \dots, \|v_N\|$. Then, for $j \geq N$ we have, using the triangle inequality,

$$\|v_j\| = \|v_j - v_N + v_N\| \leq \|v_j - v_N\| + \|v_N\| < 1 + M_N,$$

giving the result by taking $M = M_N + 1$. ■

Since we often deal simultaneously with seminorms rather than just norms, it is useful to record what is different about the two cases. What we lose for seminorms is the uniqueness of limits for convergent sequences.

6.2.5 Proposition ((Non)uniqueness of limits for (semi)normed vector spaces) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a seminormed $\mathbb{F}$-vector space. If a sequence $(\mathrm{v}_j)_{j\in\mathbb{Z}_{>0}}$ converges to limits u_0 and v_0, then*

$$\mathrm{u}_0 - \mathrm{v}_0 \in \mathsf{V}_0 \triangleq \{\mathrm{v} \in \mathsf{V} \mid \|\mathrm{v}\| = 0\}.$$

In particular, if $\|\cdot\|$ is a norm then convergent sequences have unique limits.

Proof Suppose that the sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ converges to u_0 and v_0 and let $\epsilon \in \mathbb{R}_{>0}$. Choose $N \in \mathbb{Z}_{>0}$ such that

$$\|u_0 - v_j\| \leq \tfrac{\epsilon}{2}, \quad \|v_0 - v_j\| < \tfrac{\epsilon}{2}, \qquad j \geq N.$$

For $j \geq N$ we then have

$$\|u_0 - v_0\| = \|u_0 - v_j - (v_0 - v_j)\| \leq \|u_0 - v_j\| + \|v_0 - v_j\| \leq \epsilon.$$

Therefore, $\|u_0 - v_0\| = 0$, giving the result. ■

As is the case in our previous discussions of sequences in $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, one can wonder whether all Cauchy sequences converge. In cases where they do, we call the normed vector space complete (see Definition 6.3.2). In Section 6.3 we shall see that all finite-dimensional normed vector spaces are complete (Theorem 6.3.3) but that there are easy examples of infinite-dimensional normed vector spaces that are not complete (Example 6.3.1). This is one of the factors that tends to make the theory of infinite-dimensional normed vector spaces significantly more complicated than the finite-dimensional theory. For sequences in $\mathbb{R}$ and $\mathbb{C}$ there are useful tests for convergence. There are no significant analogues for sequences in normed vector spaces.

6.2.2 Algebraic operations on sequences

Convergence is compatible with the standard algebraic operations on vector spaces.

6.2.6 Proposition (Algebraic operations on sequences) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a (semi)normed* $\mathbb{F}$*-vector space. Let* $(\mathrm{u_j})_{\mathrm{j}\in\mathbb{Z}}$ *and* $(\mathrm{v_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be sequences in* V *converging to* $\mathrm{u_0}$ *and* $\mathrm{v_0}$*, respectively, let* $(\mathrm{a_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *be a sequence in* $\mathbb{F}$ *converging to* $\mathrm{a_0}$*, and let* $\mathrm{a} \in \mathbb{F}$*. Then the following statements hold:*

(i) the sequence $(\mathrm{av_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* $\mathrm{av_0}$*;*

(ii) the sequence $(\mathrm{u_j} + \mathrm{v_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* $\mathrm{u_0} + \mathrm{v_0}$*;*

(iii) the sequence $(\mathrm{a_j v_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* $\mathrm{a_0 v_0}$*.*

Proof (i) The result is trivially true for $a = 0$, so let us suppose that $a \neq 0$. Let $\epsilon > 0$ and choose $N \in \mathbb{Z}_{>0}$ such that $\|v_j - v_0\| < \frac{\epsilon}{|a|}$. Then, for $j \geq N$,

$$\|av_j - av_0\| = |a|\|v_j - v_0\| < \epsilon.$$

(ii) Let $\epsilon > 0$ and take $N_1, N_2 \in \mathbb{Z}_{>0}$ such that

$$\|u_j - u_0\| < \tfrac{\epsilon}{2}, \quad j \geq N_1, \qquad \|v_j - v_0\| < \tfrac{\epsilon}{2}, \quad j \geq N_2.$$

Then, for $j \geq \max\{N_1, N_2\}$,

$$\|u_j + v_j - (u_0 + v_0)\| \leq \|u_j - u_0\| + \|v_j - v_0\| = \epsilon,$$

using the triangle inequality.

(iii) Let $\epsilon > 0$ and define $N_1, N_2, N_3 \in \mathbb{Z}_{>0}$ such that

$$\begin{aligned}
&|a_j - a_0| < 1, \qquad j \geq N_1, \quad \implies \quad |a_j| < |a_0| + 1, \qquad j \geq N_1,\\
&|a_j - a_0| < \frac{\epsilon}{2(|a_0| + 1)}, \qquad j \geq N_2,\\
&\|v_j - v_0\| < \frac{\epsilon}{2(\|v_0\| + 1)}, \qquad j \geq N_2.
\end{aligned}$$

Then, for $j \geq \max\{N_1, N_2, N_3\}$,

$$\begin{aligned}
\|a_j v_j - a_0 v_0\| &= \|a_j v_j - a_j v_0 + a_j v_0 - a_0 v_0\|\\
&= \|a_j(v_j - v_0) + (a_j - a_0)v_0\|\\
&\leq |a_j|\|v_j - v_0\| + |a_j - a_0|\|v_0\|\\
&\leq (|a_0| + 1)\frac{\epsilon}{2(|a_0| + 1)} + \frac{\epsilon}{2(\|v_0\| + 1)}(\|v_0\| + 1) = \epsilon,
\end{aligned}$$

as desired. ■

6.2.3 Multiple sequences

Finally, let us introduce the notion of a double sequence in a normed vector space.

6.2.7 Definition (Double sequence) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be an $\mathbb{F}$-vector space. A ***double sequence*** in V is a family of elements of V indexed by $\mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$. We denote a double sequence by $(v_{jk})_{j,k \in \mathbb{Z}_{>0}}$, where v_{jk} is the image of $(j,k) \in \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$ in V. •

For double sequences we have the following notions of convergence.

6.2.8 Definition (Convergence of double sequences) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|$ be a (semi)normed $\mathbb{F}$-vector space, and let $v_0 \in \mathsf{V}$. A double sequence $(v_{jk})_{j,k \in \mathbb{Z}_{>0}}$:

(i) ***converges to*** $\mathbf{v_0}$, and we write $\lim_{j,k\to\infty} v_{jk} = v_0$, if, for each $\epsilon > 0$, there exists $N \in \mathbb{Z}_{>0}$ such that $\|v_0 - v_{jk}\| < \epsilon$ for $j, k \geq N$;

(ii) has v_0 as a ***limit*** if it converges to v_0.

(iii) is ***convergent*** if it converges to some member of V;

(iv) ***diverges*** if it does not converge. •

missing stuff

Exercises

6.2.1 In the $\mathbb{F}$-vector space $\mathbb{F}_0^\infty$, if possible find sequences with the following properties:

(a) Cauchy in the ∞-norm but not the 2-norm;

(b) Cauchy in the 2-norm but not the 1-norm;

(c) Cauchy in the 1-norm;

(d) Cauchy in the 1-norm but not the 2-norm;

(e) Cauchy in the 2-norm but not the ∞-norm.

6.2.2 Give an example of a sequence in $\mathsf{C}^0([0,1];\mathbb{R})$ that is Cauchy with respect to the norm $\|\cdot\|_1$ but not with respect to the norm $\|\cdot\|_2$.
Hint: *Consider the function* $\mathrm{f}\colon [0,1] \to \mathbb{R}$ *defined by*

$$\mathrm{f}(\mathrm{x}) = \begin{cases} \mathrm{x}^{-1/2}, & \mathrm{x} \in (0,1], \\ 0, & \mathrm{x} = 0. \end{cases}$$

6.2.3 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. Let $(u_j)_{j\in\mathbb{Z}_{>0}}$ and $(v_j)_{j\in\mathbb{Z}_{>0}}$ be Cauchy sequences in V, let $(a_j)_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathbb{F}$, and let $a \in \mathbb{F}$.

(a) Show that $(av_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence.

(b) Show that $(a_j v_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence.

(c) Show that $(u_j + v_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence.

Section 6.3

Completeness and completions

In Theorem 2.3.5 we showed that the set of real numbers is complete in that every Cauchy sequence of real numbers converges. In Theorem **??** we used the completeness of $\mathbb{R}$ to conclude that $\mathbb{R}^n$ is complete. As we shall see in Theorem 6.3.3, every finite-dimensional normed vector space is complete. This is not true for infinite-dimensional normed vector spaces, and so for these spaces the notion of completeness has teeth: in infinite-dimensional normed vector spaces there may well be Cauchy sequences that do not converge.

For reasons that are may not be perfectly clear initially, completeness is an essential property for a normed vector space to possess. If one is confronted with a normed vector space that is not complete, the first thing one does is complete it. We have already seen in Section **??** how this works for metric spaces, and the same ideas apply for normed vector spaces. Completions are easier to understand in general than they are in specific cases. This will become painfully clear in some of the examples in Section 6.7.

Do I need to read this section? Completeness is important, so the basic ideas in this section should be understood. The technicalities can be glossed over on a first reading. •

6.3.1 Completeness (Banach spaces)

Let us begin with two examples that illustrate that for normed vector spaces, the notions of Cauchy sequences and convergent sequences are not the same.

6.3.1 Examples (Nonconvergent Cauchy sequences)

1. First consider the normed vector space $(\mathbb{F}_0^\infty, \|\cdot\|_1)$ of Example 6.1.3–6. Consider the sequence $(s_k)_{k\in\mathbb{Z}_{>0}}$ in $\mathbb{F}_0^\infty$ defined by asking that s_k be the sequence $(v_{kj})_{j\in\mathbb{Z}_{>0}}$ with

$$v_{kj} = \begin{cases} \frac{1}{j^2}, & j \in \{1,\dots,k\}, \\ 0, & j > k. \end{cases}$$

Thus the sequence s_k is the truncation to k terms of the sequence $(\frac{1}{j^2})_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$. We claim that this is a Cauchy sequence. Indeed, let $\epsilon > 0$ and choose $N \in \mathbb{Z}_{>0}$ sufficiently large that, for $k, l \geq N$ with $l > k$,

$$\sum_{j=k+1}^{l} \frac{1}{j^2} < \epsilon.$$

This is possible since the series $\sum_{j=1}^\infty \frac{1}{j^2}$ is convergent by Example 2.4.2–**??**, and so its sequence of partial sums is Cauchy. Now let $k, l \geq N$ with $l > k$ and compute

$$\|s_l - s_k\|_1 = \sum_{j=1}^{\infty} |v_{lj} - v_{kj}| = \sum_{j=k+1}^{l} \frac{1}{j^2} < \epsilon.$$

Thus the sequence $(s_k)_{k\in\mathbb{Z}_{>0}}$ is indeed Cauchy. However, it does not converge, as we now show. Suppose that $\sigma = (v_j)_{j\in\mathbb{Z}_{>0}}$ is an element of $\mathbb{F}_0^\infty$ such that $\lim_{k\to\infty}\|\sigma - s_k\|_1 = 0$, i.e., such that $(s_k)_{k\in\mathbb{Z}_{>0}}$ converges to σ. We claim that this implies that $v_j = \frac{1}{j^2}$ for each $j \in \mathbb{Z}_{>0}$. Indeed, suppose that $v_{j_0} \neq \frac{1}{j_0^2}$ for some $j_0 \in \mathbb{Z}_{>0}$. Then

$$\|\sigma - s_k\|_1 = \sum_{j=1}^{\infty} |v_j - s_k| \geq \left|v_{j_0} - \tfrac{1}{j_0^2}\right|$$

for every $k \in \mathbb{Z}_{>0}$. This implies that if $v_{j_0} \neq \frac{1}{j_0^2}$ for some $j_0 \in \mathbb{Z}_{>0}$ then $(s_k)_{k\in\mathbb{Z}_{>0}}$ cannot converge to σ. However, the sequence $(\frac{1}{j^2})_{j\in\mathbb{Z}_{>0}}$ is not in $\mathbb{F}_0^\infty$, as so we conclude that the sequence $(s_k)_{k\in\mathbb{Z}_{>0}}$ does not converge.

2. We work next with the normed vector space $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$ of Example 6.1.3–9. In this vector space, consider the sequence of functions $(f_j)_{j\in\mathbb{Z}_{>0}}$ given by

$$f_j(x) = \begin{cases} 0, & x \in [0, \frac{1}{2} - \frac{1}{2j}], \\ 2jx + 1 - j, & x \in (\frac{1}{2} - \frac{1}{2j}, \frac{1}{2}), \\ 1, & x \in [\frac{1}{2}, 1]. \end{cases}$$

In Figure 6.1 a few terms in this sequence are graphed. Suppose that $k \geq j$ so that the function $f_j - f_k$ is positive. A simple computation gives

$$\begin{aligned} \|f_j - f_k\|_1 &= \int_0^1 |f_j(x) - f_k(x)|\,\mathrm{d}x \\ &= \int_0^1 (f_j(x) - f_k(x))\,\mathrm{d}x \\ &= \int_0^1 f_j(x)\,\mathrm{d}x - \int_0^1 f_k(x)\,\mathrm{d}x \\ &= \frac{1}{2} + \frac{1}{4j} - \frac{1}{2} - \frac{1}{4k} = \frac{1}{4j} - \frac{1}{4k}. \end{aligned}$$

Now let $\epsilon > 0$ and take $N = \lceil \frac{1}{2\epsilon} \rceil$. This means that for any $j \geq N$ we have

$$j \geq N \geq \frac{1}{2\epsilon} \quad \Longrightarrow \quad \frac{1}{2j} \leq \epsilon.$$

We then have, for $j, k \geq N$,

$$\|f_j - f_k\|_1 = \left|\frac{1}{4j} - \frac{1}{4k}\right| < \frac{1}{4j} + \frac{1}{4k} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

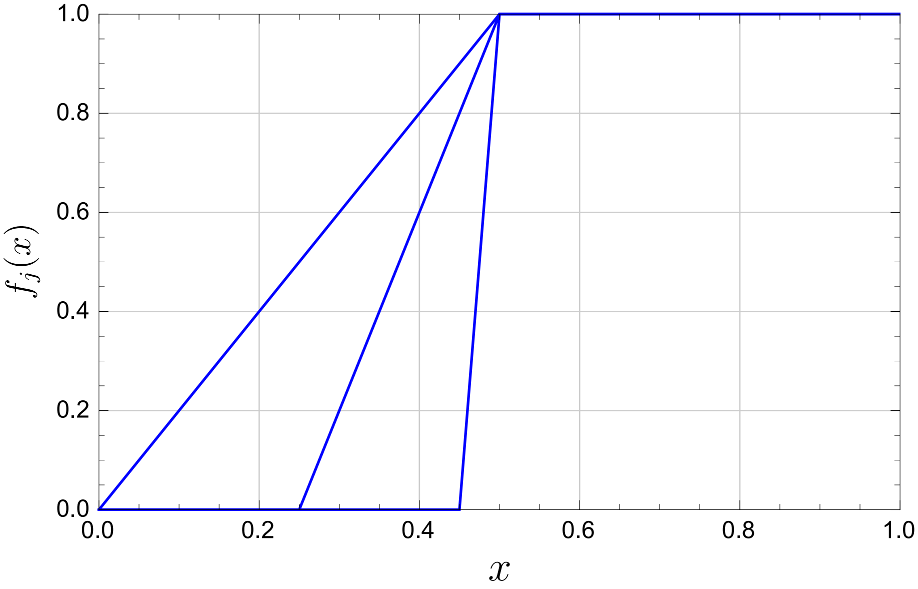


Figure 6.1 A Cauchy sequence (f_1, f_2, and f_{10} are shown) in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$

This shows that $(f_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. However, it is evident that for any $x \in [0,1]$ we have

$$\lim_{j\to\infty} f_j(x) = f(x)$$

where $f\colon [0,1] \to \mathbb{R}$ is the function

$$f(x) = \begin{cases} 0, & x \in [0, \frac{1}{2}), \\ 1, & x \in [\frac{1}{2}, 1]. \end{cases}$$

Note that $f \notin \mathsf{C}^0([0,1];\mathbb{R})$. One might want to conclude that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ does not converge since it converges pointwise to a discontinuous function. However, we should not really feel so comfortable with our knowledge of the normed vector space $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$ at this point. Thus we prove a lemma that really settles that $(f_j)_{j\in\mathbb{Z}_{>0}}$ does not, in fact, converge.

1 **Lemma** *Consider the sequence* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *as above. If* $\mathrm{g} \in \mathsf{C}^0([0,1];\mathbb{R})$ *is such that the sequence* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* g *in the norm* $\|\cdot\|_1$*, then*

$$\mathrm{g}(\mathrm{x}) = \begin{cases} 0, & \mathrm{x} \in (0, \frac{1}{2}), \\ 1, & \mathrm{x} \in (\frac{1}{2}, 1). \end{cases}$$

Proof Suppose that $g(x_0) > 0$ for some $x_0 \in [0, \frac{1}{2})$. Then, by continuity of g, there exists $\delta \in \mathbb{R}_{>0}$ such that

$$(x_0 - \delta, x_0 + \delta) \subseteq (0, \tfrac{1}{2})$$

and such that $g(x) > \frac{1}{2}g(x_0)$ for all $x \in (x_0 - \delta, x_0 + \delta)$ has the same sign as $g(x_0)$. Let $N \in \mathbb{Z}_{>0}$ be sufficiently large that $f_N(x) = 0$ for all $x \in (x_0 - \delta, x_0 + \delta)$. It then holds that for $j \geq N$ we have

$$\begin{aligned} \|g - f_j\|_1 &= \int_0^1 |g(x) - f_j(x)|\,\mathrm{d}x \geq \int_{x_0-\delta}^{x_0+\delta} |g(x) - f_j(x)|\,\mathrm{d}x \\ &= \int_{x_0-\delta}^{x_0+\delta} |g(x)|\,\mathrm{d}x \geq \delta g(x_0). \end{aligned}$$

This shows that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ cannot converge to g if $g(x_0) > 0$ for some $x_0 \in (0, \frac{1}{2})$. A completely similar argument shows that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ cannot converge to g if $g(x_0) < 0$ for some $x_0 \in (0, \frac{1}{2})$.

Now suppose that $g(x_0) > 1$ for some $x_0(\frac{1}{2}, 1)$. Then there exists $\delta > 0$ such that

$$(x_0 - \delta, x_0 + \delta) \subseteq (\tfrac{1}{2}, 1)$$

and such that $g(x) - 1 > \frac{1}{2}(g(x_0) - 1)$ for all $x \in (x_0 - \delta, x_0 + \delta)$. Then, for any $j \in \mathbb{Z}_{>0}$,

$$\begin{aligned} \|g - f_j\|_1 &= \int_0^1 |g(x) - f_j(x)|\,\mathrm{d}x \geq \int_{x_0-\delta}^{x_0+\delta} |g(x) - f_j(x)|\,\mathrm{d}x \\ &= \int_{x_0-\delta}^{x_0+\delta} |g(x) - 1|\,\mathrm{d}x \geq \delta(g(x_0) - 1). \end{aligned}$$

This shows that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ cannot converge to g if $g(x_0) > 1$ for some $x_0 \in (\frac{1}{2}, 1)$. A completely similar argument shows that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ cannot converge to g if $g(x_0) < 1$ for some $x_0 \in (\frac{1}{2}, 1)$. ▼

There is obviously no continuous function satisfying the conditions of the lemma. Thus we have found a Cauchy sequence in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$ that does not converge. •

The examples show something very important: that there is a genuine distinction between Cauchy sequences and convergent sequences. Moreover, normed vector spaces where the two notions agree are important.

6.3.2 Definition (Completeness, Banach space) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. A normed $\mathbb{F}$-vector space $(\mathsf{V}, \|\cdot\|)$ is ***complete*** if every Cauchy sequence in V converges. A ***$\mathbb{F}$-Banach space*** is a complete normed $\mathbb{F}$-vector space. •

The following result is important in the same way that completeness of $\mathbb{R}$ is important.

6.3.3 Theorem (Completeness of finite-dimensional normed vector spaces) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $(\mathsf{V}, \|\cdot\|)$ *is a finite-dimensional normed* $\mathbb{F}$*-vector space, then* V *is complete.*

Proof Let $\{e_1, \dots, e_n\}$ be a basis for V which defines an isomorphism $\iota \colon \mathsf{V} \to \mathbb{F}^n$ by

$$\iota(v_1 e_1 + \dots + v_n e_n) = (v_1, \dots, v_n).$$

Define a norm $\|\cdot\|_2$ on V by $\|v\|_2 = \|\iota(v)\|_2$ where $\|\cdot\|_2$ also denotes the standard norm on $\mathbb{F}^n$. This is a norm, cf. the proof of Proposition 6.1.4. By Theorem 6.1.15 it follows that there exists $C \in \mathbb{R}_{>0}$ such that

$$C^{-1}\|v\|_2 \le \|v\| \le C\|v\|_2.$$

Now let $(v_j)_{j \in \mathbb{Z}_{>0}}$ be a Cauchy sequence in V. Let's write

$$v_j = v_{j1} e_1 + \dots + v_{jn} e_n$$

for $v_{jl} \in \mathbb{F}$, $j \in \mathbb{Z}_{>0}$, $l \in \{1, \dots, n\}$. For $\epsilon \in \mathbb{R}_{>0}$ let $N \in \mathbb{Z}_{>0}$ by such that $\|v_j - v_k\| < C^{-1}\epsilon$ for $j, k \in \mathbb{Z}_{>0}$. We then have

$$C^{-1}\epsilon > \|v_j - v_k\| \ge C^{-1}\|v_j - v_k\|_2 = C^{-1}\Big(\sum_{l=1}^{n} |v_{jl} - v_{kl}|\Big)^{1/2} \ge C^{-1}|v_{jl_0} - v_{kl_0}|$$

for $j, k \ge N$ and for each $l_0 \in \{1, \dots, n\}$. Thus $|v_{jl_0} - v_{kl_0}| < \epsilon$ for $j, k \ge N$ and for each $l_0 \in \{1, \dots, n\}$. Thus $(v_{jl_0})_{j \in \mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{F}$ for each $l_0 \in \{1, \dots, n\}$. Since $\mathbb{F}$ is complete by Theorem **??** it follows that there exists $v_{l_0} \in \mathbb{F}$, $l_0 \in \{1, \dots, n\}$, such that $\lim_{j \to \infty} v_{jl_0} v_{l_0}$. Now define $v = v_1 e_1 + \dots + v_n e_n$. We claim that $(v_j)_{j \in \mathbb{Z}_{>0}}$ converges to v. To see this, for $\epsilon \in \mathbb{R}_{>0}$ let $N \in \mathbb{Z}_{>0}$ be such that $\|v_l - v_{jl}\| < \frac{\epsilon}{C\sqrt{n}}$. Then

$$\|v - v_j\| \le C\|v - v_j\|_2 = C\Big(\sum_{l=1}^{n} |v_l - v_{jl}|^2\Big)^{1/2} \le C\Big(\sum_{l=1}^{n} \Big(\frac{\epsilon}{C\sqrt{n}}\Big)^2\Big)^{1/2} = \epsilon,$$

as desired. ■

6.3.2 Why completeness is important

We have now seen completeness arise in three important cases. The first was with the incompleteness of the rational numbers and the second and third were in Example 6.3.1. It is fair to ask, "Who cares whether a normed vector space is complete?" In this section we address this.

First let us consider the simple case of the incompleteness of the rational numbers. Rational numbers are fairly simple to define and pretty easy to understand. Real numbers are somewhat more difficult to define, and we think we understand them only because we live in a world where the notion of a real number has been accepted for so long that they are as integral a part of science as are the integers. However, it is worth reflecting that the notion of numbers that were not rational numbers has not always been as acceptable as it is now. Indeed, the development of mathematics is marked by strong resistance to any of the "unusual" kinds of new numbers that arose, whether they be negative numbers, real numbers, or complex numbers. As concerns real numbers, many Greek mathematicians were dedicated to the existence only of rational numbers. There is an amusing story—completely unsubstantiated by any historical record and thus almost certainly false—that a student of Pythagoras was thrown into the sea for proving that $\sqrt{2}$ was not rational. It is also worth reflecting that, if one is only interested in computation, rational numbers are all one can represent in a digital computer. Thus it is difficult to justify the construction of the real numbers from a purely practical point of view. So why are the real numbers important? They are important precisely because they are complete. It is completeness that makes true "obvious" statements like, "every bounded increasing sequence converges." Relatively simple ideas like continuity and differentiability of functions, the Riemann integral, convergence of sequences of functions, all rely on the completeness of the real numbers for their power. Scientific life would be very difficult and complicated without the completeness of the real numbers.

The point of the above paragraph is this:

1. The real numbers arise in a natural way from the incompleteness of the rational numbers.
2. The completeness of the real numbers is not important for the purposes of computation.
3. The completeness of the real numbers is important for the very basic ideas we use every day concerning real variables and functions of a real variable.
4. You are probably comfortable with the real numbers, but this is only because of societal norms.

Now let us think about the notion of completeness in normed vector spaces. Indeed, let us think specifically about Example 6.3.1–2. In that example we saw that there is a simple Cauchy sequence in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$ that does not converge. But the sequence of functions certainly converges to a perfectly nice, albeit discontinuous, function. So why not just include this limit function in our set and move on? Well, one can certainly do this, but it also leads to the question, "What are the functions that we need to add to $\mathsf{C}^0([0,1];\mathbb{R})$ in order to be sure that all Cauchy sequences of continuous functions converge?" This is a little like saying that, since $\sqrt{2}$ is irrational, why not just add it to our collection of numbers and move on (the result would be the field extension $\mathbb{Q}(\sqrt{2})$). One could do this, but then eventually one would need to address the matter of what other kinds of numbers need to be added to the rational numbers. Thinking about things in this sort of *ad hoc* way is not satisfying, and is really just faking your way around the real issue, which is this: *one should be sure to always be dealing with complete normed vector spaces*.

The difficulty that arises, as we shall see in Section 6.7.7, is that it is difficult to describe the set of functions that need to be added to $\mathsf{C}^0([0,1];\mathbb{R})$ in order to ensure completeness with respect to the norm $\|\cdot\|_1$. But the point is that just because it is difficult does not mean that it is not important to do. It *is* important to do. Indeed, at some point one *must* do it.

6.3.3 Completeness and direct sums and quotients

In this section we consider how completeness interacts with direct sums and quotients. We first consider direct sums. Recall from Theorem 6.1.17 that if $((\mathsf{V}_i, \|\cdot\|_i))_{i\in I}$ is a family of normed vector spaces then we define a norm $\|\cdot\|_I$ on the direct sum $\bigoplus_{i\in I} \mathsf{V}_i$ by

$$\|\phi\|_I = \sum_{i\in I} \|\phi(i)\|_i,$$

the sum making sense since it is finite.

6.3.4 Proposition (Completeness of direct sums of Banach spaces) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $((\mathsf{V}_{\mathrm{i}}, \|\cdot\|_{\mathrm{i}}))_{\mathrm{i}\in\mathrm{I}}$ be a family of $\mathbb{F}$-Banach spaces. Then $(\bigoplus_{\mathrm{i}\in\mathrm{I}} \mathsf{V}_{\mathrm{i}}, \|\cdot\|_{\mathrm{I}})$ is complete if and only if I is finite.*

Proof First suppose that I is finite and so take $I = \{1, \dots, k\}$. Let us denote elements of $\bigoplus_{l=1}^{k} \mathsf{V}_j$ as $(v_1, \dots, v_k)$. Let $((v_{1j}, \dots, v_{kj}))_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\bigoplus_{l=1}^{k} \mathsf{V}_j$. We claim that, for each $l \in \{1, \dots, k\}$, $(v_{lj})_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in V_l. Let $\epsilon \in \mathbb{R}_{>0}$ and take $N \in \mathbb{Z}_{>0}$ sufficiently large that

$$\|(v_{1j}, \dots, v_{kj}) - (v_{1m}, \dots, v_{km})\|_I < \epsilon, \qquad j, m \geq N.$$

Since

$$\|(v_{1j}, \dots, v_{kj}) - (v_{1m}, \dots, v_{km})\|_I = \|v_{1j} - v_{1m}\|_1 + \dots + \|v_{kj} - v_{km}\|_k$$

it follows that

$$\|v_{lj} - v_{lm}\|_l < \epsilon, \qquad j, m \geq N,$$

and so the sequence $(v_{lj})_{j\in\mathbb{Z}_{>0}}$ is indeed Cauchy. Therefore, since V_l is a Banach space, the sequence converges to $v_l \in \mathsf{V}_l$. We next claim that the sequence $((v_{1j}, \dots, v_{kj}))_{j\in\mathbb{Z}_{>0}}$ converges to $(v_1, \dots, v_k)$. Indeed, let $\epsilon \in \mathbb{R}_{>0}$ and take $N \in \mathbb{Z}_{>0}$ sufficiently large that

$$\|v_{lj} - v_l\|_l < \tfrac{\epsilon}{k}, \qquad l \in \{1, \dots, k\},\ j \geq N.$$

Then

$$\|(v_{1j}, \dots, v_{kj}) - (v_1, \dots, v_k)\|_I = \|v_{1j} - v_1\|_1 + \dots + \|v_{kj} - v_k\|_k < \epsilon,$$

for $j \geq N$, giving the desired convergence.

Next suppose that I is infinite and, for each $i \in I$, choose $v_i \in \mathsf{V}_i$ such that $\|v_i\|_i = 1$. Let $\{i_l\}_{l\in\mathbb{Z}_{>0}}$ be a set of distinct elements of I and then define a sequence $(\phi_k)_{k\in\mathbb{Z}_{>0}}$ in $\bigoplus_{i\in I} \mathsf{V}_i$ by

$$\phi_k(i) = \begin{cases} 2^{-j} v_{i_j}, & i = i_j,\ j \in \{1, \dots, k\}, \\ 0, & \text{otherwise.} \end{cases}$$

We claim that $(\phi_k)_{k\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. Indeed, let $\epsilon > 0$ and let $N \in \mathbb{Z}_{>0}$ be such that for $k, m \geq N$ with $m > k$ we have $\sum_{j=k+1}^{m} < \epsilon$. This is possible since the series $\sum_{j=1}^{\infty} 2^{-j}$ converges by Example 2.4.2–??. Now note that, for $k, m \geq N$ with $m > k$ we have

$$\|\phi_k - \phi_m\|_I = \sum_{j=k+1}^{m} \|2^{-j} v_{i_j}\|_{i_j} = \sum_{j=k+1}^{m} 2^{-j} < \epsilon,$$

showing that the sequence $(\phi_k)_{k\in\mathbb{Z}_{>0}}$ is indeed Cauchy. However, the sequence does not converge. Indeed, if $\phi \in \bigoplus_{i\in I} \mathsf{V}_i$ has the property that $\lim_{k\to\infty} \|\phi - \phi_k\|_I = 0$ then this implies that $\phi(i_j) = 2^{-j} v_{i_j}$ for $j \in \mathbb{Z}_{>0}$, cf. Example 6.3.1–1. But then $\phi \notin \bigoplus_{i\in I} \mathsf{V}_i$. ■

In Section 6.7.3 we will revisit the matter of the completeness of direct sums.

For now we turn to quotients. We recall from Proposition 6.1.19 the definition of the norm $\|\cdot\|_{/\mathsf{U}}$ on V/U.

6.3.5 Proposition (Quotients of Banach spaces by closed subspaces are Banach spaces) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, if $(\mathsf{V}, \|\cdot\|)$ is an $\mathbb{F}$-Banach space, and if U is a closed subspace of V, then $(\mathsf{V}/\mathsf{U}, \|\cdot\|_{/\mathsf{U}})$ is an $\mathbb{F}$-Banach space.*

Proof We already know from Proposition 6.1.19 that $(\mathsf{V}/\mathsf{U}, \|\cdot\|_{/\mathsf{U}})$ is a normed vector space, so it is completeness that e must prove here. Let $(v_j + \mathsf{U})_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence. By passing to a subsequence if necessary we can suppose that $\|(v_{j+1} - v_j) + \mathsf{U}\|_{/\mathsf{U}} < 2^{-j}$, $j \in \mathbb{Z}_{>0}$. By definition of $\|\cdot\|_{/\mathsf{U}}$ this means that there exists $u_2 \in \mathsf{U}$ such that $\|v_2 + u_2 - v_1\| < 2^{-1}$. Define $v_2' = v_2 + u_2$. Similarly, there exists $u_3 \in \mathsf{U}$ such that $\|v_3 + u_3 - v_2\| < 2^{-2}$. Define $v_3' = v_3 + u_3$. Proceeding in this way we define a sequence $(v_j')_{j\in\mathbb{Z}_{>0}}$ such that $\|v_{j+1}' - v_j'\| < 2^{-j}$ and such that $v_j' + \mathsf{U} = v_j + \mathsf{U}$ for $j \in \mathbb{Z}_{>0}$. In particular, the sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ is Cauchy and so converges to some $v \in \mathsf{V}$ since V is complete. Then, by Theorem 6.5.2,

$$\lim_{j\to\infty}(v_j' + \mathsf{U}) = (\lim_{j\to\infty} v_j') + \mathsf{U} = v + \mathsf{U}$$

since the projection from V to V/U is continuous. ■

6.3.4 Completions

Having been confronted in Section 6.3.1 with the reality of normed vector spaces that are not complete, and having seen evidence of the importance of completeness in Section 6.3.2, it becomes important to know the answer to this question: "What do we do when we have an incomplete normed vector space?" The answer is: "We complete it!"

The notion of a completion was discussed in detail in Section **??** for metric spaces. Since normed vector spaces are metric spaces by Proposition 6.1.7, that entire discussion can be transported here to define the completion of a normed vector space. However, we will develop at least some of this discussion independently.

The main result is the following. In the statement of the result we make reference to the notion of an isomorphism of normed vector spaces. We will not formally get to this idea until Section 6.5.2, but let us just say here that an isomorphism of normed vector spaces is an invertible linear map that is continuous and has a continuous inverse.

6.3.6 Theorem (Completion of a normed vector space) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. Then there exists a Banach space $(\overline{\mathsf{V}}, \overline{\|\cdot\|})$ with the following properties:*

(i) there exists an injective linear map $\iota_{\mathsf{V}}\colon \mathsf{V} \to \overline{\mathsf{V}}$ such that $\overline{\|\iota_{\mathsf{V}}(\mathrm{v})\|} = \|\mathrm{v}\|$ for every $\mathrm{v} \in \mathsf{V}$;

(ii) for each $\overline{\mathrm{v}} \in \overline{\mathsf{V}}$ there exists a sequence $(\mathrm{v_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ in V such that $(\iota_{\mathsf{V}}(\mathrm{v_j}))_{\mathrm{j}\in\mathbb{Z}_{>0}}$ converges to $\overline{\mathrm{v}}$.

Such a Banach space $(\overline{\mathsf{V}}, \overline{\|\cdot\|})$ *is a* ***completion*** *of* $(\mathsf{V}, \|\cdot\|)$.

Furthermore, if $(\overline{\mathsf{V}}_1, \overline{\|\cdot\|_1})$ *and* $(\overline{\mathsf{V}}_2, \overline{\|\cdot\|_2})$ *are two completions of* $(\mathsf{V}, \|\cdot\|)$ *with* $\iota_{\mathsf{V},1}\colon \mathsf{V} \to \overline{\mathsf{V}}_1$ *and* $\iota_{\mathsf{V},2}\colon \mathsf{V} \to \overline{\mathsf{V}}_2$ *being the corresponding injective linear maps, then there exists an isomorphism* $\mathsf{L}\colon \overline{\mathsf{V}}_1 \to \overline{\mathsf{V}}_2$ *of Banach space such that the following diagram commutes:*

$$\begin{tikzcd}
& \mathsf{V} \arrow[dl, "\iota_{\mathsf{V},1}"'] \arrow[dr, "\iota_{\mathsf{V},2}"] & \\
\overline{\mathsf{V}}_1 \arrow[rr, "\mathsf{L}"'] & & \overline{\mathsf{V}}_2
\end{tikzcd}$$

Proof Many of the details of this proof follow that of Theorem **??**, and we therefore omit them, only making reference to the existing proof.

We let $\mathrm{CS}(\mathsf{V})$ denote the collection of Cauchy sequences in V. If we define vector addition and scalar multiplication by

$$(u_j)_{j\in\mathbb{Z}_{>0}} + (v_j)_{j\in\mathbb{Z}_{>0}} = (u_j + v_j)_{j\in\mathbb{Z}_{>0}}, \quad a(v_j)_{j\in\mathbb{Z}_{>0}} = (av_j)_{j\in\mathbb{Z}_{>0}},$$

then $\mathrm{CS}(\mathsf{V})$ is an $\mathbb{F}$-vector space by Exercise 6.2.3.

For a Cauchy sequences $(v_j)_{j\in\mathbb{Z}_{>0}}$ let us define

$$\|\widetilde{(v_j)_{j\in\mathbb{Z}_{>0}}}\| = \lim_{j\to\infty} \|v_j\|.$$

To make the connection with the proof of Theorem **??** we note that we can define

$$\tilde{d}((u_j)_{j\in\mathbb{Z}_{>0}}, (v_j)_{j\in\mathbb{Z}}) = \lim_{j\to\infty} \|u_j - v_j\|.$$

Then we obviously have

$$\|\widetilde{(v_j)_{j\in\mathbb{Z}_{>0}}}\| = \tilde{d}((v_j)_{j\in\mathbb{Z}_{>0}}, (0)_{j\in\mathbb{Z}_{>0}}).$$

This identity can be used to easily prove many of the assertions we are about to make about $\widetilde{\|\cdot\|}$. In particular, the definition of $\widetilde{\|\cdot\|}$ is shown to make sense in that the limit exists. Moreover, $\widetilde{\|\cdot\|}$ is readily seen to be a seminorm on $\mathrm{CS}(\mathsf{V})$. For example, we compute

$$\|\widetilde{a(v_j)_{j\in\mathbb{Z}_{>0}}}\| = \lim_{j\to\infty} \|av_j\| = |a| \lim_{j\to\infty} \|v_j\| = |a| \|\widetilde{(v_j)_{j\in\mathbb{Z}_{>0}}}\|.$$

(Note that in the third step we make use of continuity of the norm which we will prove as Proposition 6.5.4.) The remaining seminorm properties follow just as do the corresponding assertions from Theorem **??**.

We now let $(\overline{\mathsf{V}}, \overline{\|\cdot\|})$ be the normed vector space associated with $(\mathrm{CS}(\mathsf{V}), \widetilde{\|\cdot\|})$ as in Theorem 6.1.8. Note that $(\overline{\mathsf{V}}, \overline{\|\cdot\|})$ as in Theorem 6.1.8 is the normed vector space whose associated metric space is the metric space $(\overline{\mathsf{V}}, \overline{d})$ of Theorem **??**. From Exercise 6.3.4 it immediately follows that $(\overline{\mathsf{V}}, \overline{\|\cdot\|})$ is a Banach space.

Recalling from Theorem 6.1.8 that $\overline{\mathsf{V}}$ is a quotient of $\mathrm{CS}(\mathsf{V})$ by a subspace, denote by $\pi_{\mathsf{V}}\colon \mathrm{CS}(\mathsf{V}) \to \overline{\mathsf{V}}$ the canonical projection. Now define $\iota_{\mathsf{V}}\colon \mathsf{V} \to \overline{\mathsf{V}}$ by $\iota_{\mathsf{V}}(v) = \pi_{\mathsf{V}}((v)_{j\in\mathsf{V}})$. As for the corresponding assertion from Theorem **??**, we readily show that $\overline{\|\iota_{\mathsf{V}}(v)\|} = \|v\|$

for each $v \in \mathsf{V}$. Since the injection ι_{V} of V into $\overline{\mathsf{V}}$ is the same as the injection in the proof of Theorem **??**, it follows from Theorem **??** that for any $\overline{v} \in \overline{\mathsf{V}}$ there is a sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ for which $(\iota_{\mathsf{V}}(v_j))_{j\in\mathbb{Z}_{>0}}$ converges to $\overline{v}$.

Now we prove the final assertion of the theorem, letting $(\overline{\mathsf{V}}_1, \overline{\|\cdot\|_1})$ and $(\overline{\mathsf{V}}_1, \overline{\|\cdot\|_2})$ be completions of $(\mathsf{V}, \|\cdot\|)$. Let $\overline{v}_1 \in \overline{\mathsf{V}}_1$ and let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence for which $(\iota_{\mathsf{V},1})_{j\in\mathbb{Z}_{>0}}$ converges to $\overline{v}_1$. Thus $(\iota_{\mathsf{V},1}(v_j))_{j\in\mathbb{Z}_{>0}}$ is Cauchy. Since $\iota_{\mathsf{V},1}$ preserves the norm, one easily shows that $(v_j)_{j\in\mathbb{Z}_{>0}}$ is Cauchy. Since $\iota_{\mathsf{V},2}$ also preserves the norm, the sequence $(\iota_{\mathsf{V},2}(v_j))_{j\in\mathbb{Z}_{>0}}$ is Cauchy, and so converges since $\overline{\mathsf{V}}_2$ is complete. Let $\overline{v}_2$ denote its limit. We define $\mathsf{L}\colon \overline{\mathsf{V}}_1 \to \overline{\mathsf{V}}_2$ by $\mathsf{L}(\overline{v}_1) = \overline{v}_2$, according to the preceding construction. As with the corresponding assertion in the proof of Theorem **??**, one can show that this definition is independent of the choice of sequence converging to $\overline{v}_1$. Moreover, just as in the proof of Theorem **??**, we can show that L is a bijection and an isometry. Therefore, it is continuous and has a continuous inverse.***missing stuff*** All that remains is to show that L is linear. To see this, let $\overline{u}_1, \overline{v}_1 \in \overline{\mathsf{V}}_1$ and let $a \in \mathbb{F}$. Let $(u_j)_{j\in\mathbb{Z}_{>0}}$ and $(v_j)_{j\in\mathbb{Z}_{>0}}$ be sequences in V for which $\lim_{j\to\infty} \iota_{\mathsf{V},1}(u_j) = \overline{u}_1$ and $\lim_{j\to\infty} \iota_{\mathsf{V},1}(u_j) = \overline{u}_1$. We then have

$$\mathsf{L}(a\overline{v}_1) = \lim_{j\to\infty} \iota_{\mathsf{V},2}(av_j) = a \lim_{j\to\infty} \iota_{\mathsf{V},2}(v_j) = a\mathsf{L}(\overline{v}_1)$$

and

$$\mathsf{L}(\overline{u}_1 + \overline{v}_1) = \lim_{j\to\infty} \iota_{\mathsf{V},2}(u_j + v_j) = \lim_{j\to\infty} \iota_{\mathsf{V},2}(u_j) + \lim_{j\to\infty} \iota_{\mathsf{V},2}(v_j) = \mathsf{L}(\overline{u}_1) + \mathsf{L}(\overline{v}_1),$$

where we have used the continuity properties of the norm as in Proposition 6.5.4 below. ■

The preceding theorem is nice in that the proof is constructive. The completion consists of equivalence classes of Cauchy sequences, just as was the case for the construction of $\mathbb{R}$ in Section 2.1.2. The problem is that it may not be so easy to understand what elements in the completion "look like." For example, in Example 6.3.1 we gave two instances of incomplete normed vector spaces. For the incomplete normed vector space $(\mathbb{F}_0^\infty, \|\cdot\|_1)$ it is fairly easy to understand the completion; we do this in Section 6.7.2. However, for the incomplete normed vector space $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$ the completion is harder to understand. Indeed, try to imagine what might be the set of limits of all Cauchy sequences in $\mathsf{C}^0([0,1];\mathbb{R})$. Surely these limits can be pretty complicated! And we shall see in Section 6.7.7 that to describe these limits is possible by using Lebesgue's integral that we dedicated so much effort to in Chapter 5. Indeed, many of the examples of Banach spaces in Section 6.7 are constructed as completions. The diversity of the examples in that section should, alone, convince the reader of the importance of completeness and completions.

Exercises

6.3.1 For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ show that $(\mathbb{F}_0^\infty, \|\cdot\|_\infty)$ (see Example 6.1.3–7) is not complete.

6.3.2 Consider the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ of signals in $\mathsf{C}^0([0,1];\mathbb{R})$ as defined in Example 6.3.1–2. In this exercise, use the norm $\|\cdot\|_\infty$.

(a) Show by explicit calculation that the sequence is not a Cauchy sequence.

(b) Is it possible to deduce that the sequence is not Cauchy without doing any calculations?

6.3.3 Consider the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ of functions in $\mathsf{C}^0([0,1];\mathbb{R})$ defined by $f_j(x) = x^j$. For the vector space $\mathsf{C}^0([0,1];\mathbb{R})$ consider two norms, $\|\cdot\|_\infty$ and $\|\cdot\|_1$, defined by:

$$\|f\|_\infty = \sup\{|f(x)| \mid x\in[0,1]\},$$
$$\|f\|_1 = \int_0^1 |f(x)|\,\mathrm{d}x.$$

Answer the following questions.

(a) Sketch the graphs of the first few functions in the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$.

(b) Is the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ a Cauchy sequence in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_\infty)$?

(c) Is the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ a Cauchy sequence in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$?

(d) Does the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ converge in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_\infty)$?

(e) If the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ does not converge in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_\infty)$, does it converge in the completion of $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_\infty)$? If so, to what function does it converge?

(f) Does the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ converge in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$?

(g) If the sequence $\{f_j\}_{j\in\mathbb{Z}_{>0}}$ does not converge in $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$, does it converge in the completion of $(\mathsf{C}^0([0,1];\mathbb{R}), \|\cdot\|_1)$? If so, to what function does it converge?

6.3.4 Show that a normed vector space $(\mathsf{V}, \|\cdot\|)$ is complete if and if the associated metric space (from Proposition 6.1.7) is complete.

6.3.5 Let $((\mathsf{V}_i, \|\cdot\|_i))_{i\in I}$ be a family of normed vector spaces with $(\bigoplus_{i\in I}\mathsf{V}_i, \|\cdot\|_I)$ the corresponding direct sum normed vector space. Show that, if $(\bigoplus_{i\in I}\mathsf{V}_i, \|\cdot\|_I)$ is complete, then $(\mathsf{V}_i, \|\cdot\|_i)$ is complete for each $i\in I$.

6.3.6 Let $((\mathsf{V}_i, \|\cdot\|_i))_{i\in I}$ be a family of Banach spaces and define the norm $\|\cdot\|_{I,\infty}$ on $\bigoplus_{i\in I}\mathsf{V}_i$ by

$$\|\phi\|_{I,\infty} = \max\{|\phi(i)| \mid i\in I\}.$$

Show that $(\bigoplus_{i\in I}\mathsf{V}_i, \|\cdot\|_{I,\infty})$ is incomplete if I is infinite.

6.3.7 On the vector space $\mathsf{AC}([a,c];\mathbb{F})$ of $\mathbb{F}$-valued absolutely continuous functions on $[a,b]$, define the function $f\mapsto \|f\|$ by

$$\|f\| = \int_a^b |f(x)|\,\mathrm{d}x.$$

Answer the following questions.

(a) Show that $(\mathsf{AC}([a,b];\mathbb{F}), \|\cdot\|)$ is a normed vector space.

(b) Show that you understand why $(\mathsf{AC}([a,b];\mathbb{F}), \|\cdot\|)$ is not a Banach space by providing a nonconvergent Cauchy sequence.

Section 6.4

Series in normed vector spaces

We now consider series in normed vector spaces. While some of the development here bears a strong resemblance to that for series in $\mathbb{R}$ given in Section 2.4, there are some significant differences. In particular, we introduce two new notions of convergence, condition and unconditional convergence. The latter of these is equivalent for series in $\mathbb{R}$ to absolute convergence, as we show in Proposition 6.4.5. However, in infinite-dimensions the two notions are not equivalent, and we prove this as the nontrivial Theorem 6.4.8. Much of the rest of the development follows in the same vein as that for series in $\mathbb{R}$.

Do I need to read this section? The reader should understand the notion of a series in a normed vector space since this will be important to us in Section 7.3, which in turn is important in the theory of Fourier series. The material in Section 6.4.2, while interesting, is also somewhat technical and can be skipped at a first reading. The material in Sections 6.4.5 and 6.4.6 can likewise be overlooked until it is needed. •

6.4.1 Definitions and properties of series

A ***series*** in an $\mathbb{F}$-vector space is an expression of the form

$$\sum_{j=1}^{\infty} v_j,$$

where $v_j \in \mathsf{V}$, $j \in \mathbb{Z}_{>0}$. As with series in $\mathbb{R}$ or $\mathbb{C}$, this expression is merely symbolic (but still sensible as a formal expression) unless something can be said about its convergence. For vector spaces without any structure, series can be nothing more than formal. Fortunately, (semi)normed vector spaces have topologies defined on them, and so notions of convergence can be defined. These are as follows.

6.4.1 Definition (Convergence, absolute convergence, and conditional convergence of series) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in V and consider the series

$$S = \sum_{j=1}^{\infty} v_j.$$

The corresponding sequence of ***partial sums*** is the sequence $(S_k)_{k\in\mathbb{Z}_{>0}}$ in V defined by

$$S_k = \sum_{j=1}^{k} v_j.$$

Let $v_0 \in \mathsf{V}$. The series:

(i) is ***Cauchy*** if the sequence of partial sums is a Cauchy sequence;

(ii) ***converges to*** $\mathbf{v_0}$, and we write $\sum_{j=1}^{\infty} v_j = v_0$, if the sequence of partial sums converges to v_0;

(iii) has v_0 as a ***limit*** if it converges to v_0;

(iv) is ***convergent*** if it converges to some member of V;

(v) ***converges absolutely***, or is ***absolutely convergent***, if the series

$$\sum_{j=1}^{\infty} \|v_j\|$$

converges;

(vi) is ***unconditionally Cauchy*** if, for every bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$, the series $S_\phi = \sum_{j=1}^{\infty} v_{\phi(j)}$ is Cauchy;

(vii) ***converges unconditionally***, or is ***unconditionally convergent***, if, for every bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$, the series $S_\phi = \sum_{j=1}^{\infty} v_{\phi(j)}$ converges;

(viii) is ***conditionally Cauchy*** if it is not unconditionally Cauchy;

(ix) ***converges conditionally***, or is ***conditionally convergent***, if it is not unconditionally convergent;

(x) ***diverges*** if it does not converge. •

There are a few differences between the definitions we give here and those for given in Definition 2.4.1 for series of real numbers. These differences have real substance, so let us record why they arise.

1. In Definition 2.4.1 we did not have the notion of Cauchy series. This is because this is not necessary for series in $\mathbb{R}$ since Cauchy sequences converge. However, in infinite-dimensional normed vector spaces there may well be non-convergent Cauchy sequences. Therefore, it is useful to distinguish between Cauchy sequences of partial sums and convergent sequences of partial sums. Whenever possible we state results for Cauchy series rather than convergent series, keeping in mind that convergent series are Cauchy.
2. There is a difference between the notions of conditional convergence for series in normed vector spaces and for real numbers as given in Definition 2.4.1. There is some substance to this difference, and we shall explore this in Section 6.4.2, particularly Theorem 6.4.8.

Just as for series of real numbers and complex numbers, there is a useful relationship between the norm of a sum and the sum of the norms.

6.4.2 Proposition (Swapping summation and norm) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a (semi)normed* $\mathbb{F}$*-vector space. For a sequence* $(\mathrm{v}_j)_{j\in\mathbb{Z}_{>0}}$*, if the series* $\mathrm{S} = \sum_{j=1}^{\infty} \mathrm{v}_j$ *is absolutely convergent, then*

$$\left\|\sum_{j=1}^{\infty} \mathrm{v}_j\right\| \leq \sum_{j=1}^{\infty} \|\mathrm{v}_j\|.$$

Proof Define

$$S_m^1 = \Big\| \sum_{j=1}^m v_j \Big\|, \quad S_m^2 = \sum_{j=1}^m \|v_j\|, \qquad m \in \mathbb{Z}_{>0}.$$

By Exercise 6.4.1 we have $S_m^1 \leq S_m^2$ for each $m \in \mathbb{Z}_{>0}$. Moreover, by Proposition 6.4.5 the sequences $(S_m^1)_{m\in\mathbb{Z}_{>0}}$ and $(S_m^2)_{m\in\mathbb{Z}_{>0}}$ are Cauchy sequences in $\mathbb{R}$ and so converge. It is then clear that

$$\lim_{m\to\infty} S_m^1 \leq \lim_{m\to\infty} S_m^2,$$

which is the result. ■

While we do not have for series in normed vector spaces the bevy of tests for convergence, we do have the obvious sufficient condition.

6.4.3 Proposition (Sufficient condition for a series to diverge) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. If the sequence $(\|\mathrm{v}_\mathrm{j}\|)_{\mathrm{j}\in\mathbb{Z}_{>0}}$ does not converge to zero, then the series $\sum_{\mathrm{j}=1}^\infty \mathrm{v}_\mathrm{j}$ diverges.*

Proof Suppose that the series $\sum_{j=1}^\infty v_j$ converges to v_0 and let $(S_k)_{k\in\mathbb{Z}_{>0}}$ be the sequence of partial sums. Then $v_k = S_k - S_{k-1}$. Then

$$\lim_{k\to\infty} v_k = \lim_{k\to\infty} S_k - \lim_{k\to\infty} S_{k-1} = v_0 - v_0 = 0_{\mathsf{V}},$$

as desired. ■

6.4.2 Absolute and unconditional convergence

In this section we explore the relationship between absolute and unconditional convergence. For finite-dimensional normed vector spaces we will see that the two notions are equivalent.

Let us begin by showing why unconditional convergence is useful, in the same way we showed that absolute convergence is useful in Theorem 2.4.5.

6.4.4 Proposition (Unconditional limits are rearrangement independent) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. If the series $\sum_{\mathrm{j}=1}^\infty \mathrm{v}_\mathrm{j}$ is unconditionally convergent and converges to v_0, then, for any bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$, the series $\sum_{\mathrm{j}=1}^\infty \mathrm{v}_{\phi(\mathrm{j})}$ also converges to v_0.*

Proof In order to avoid duplication of part of the proof, we make use of the implication (ii) $\Longrightarrow$ (i) of Theorem 6.4.20. We do this in the following way. Let $S = \sum_{j=1}^\infty v_j$. Since S is unconditionally convergent it is unconditionally Cauchy by Proposition 6.2.3. By the implication (ii) $\Longrightarrow$ (i) of Theorem 6.4.20 it follows that $\sum_{j\in\mathbb{Z}_{>0}} v_j$ is Cauchy in the sense of Definition 6.4.16. Now let $\epsilon \in \mathbb{R}_{>0}$ and let $I \subseteq \mathbb{Z}_{>0}$ be a finite set with the property that

$$\Big\| \sum_{j\in J} v_j \Big\| < \frac{\epsilon}{2}$$

for any finite set J such that $J \cap I = \emptyset$. Now let $N_1 \in \mathbb{Z}_{>0}$ be such that

$$\Big\| \sum_{j=1}^k v_j - v_0 \Big\| < \frac{\epsilon}{2}$$

for every $k \geq N_1$ (this being possible since $\sum_{j=1}^{\infty} v_j$ converges to v_0) and such that $I \subseteq \{1, \ldots, N_1\}$. Let $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a bijection and choose $N_2 \in \mathbb{Z}_{>0}$ sufficiently large that $\{1, \ldots, N_1\} \subseteq \{\phi(1), \ldots, \phi(N_2)\}$. Then we write

$$\{\phi(1), \ldots, \phi(N_2)\} = \{1, \ldots, N_1\} \cup J$$

where $J \cap \{1, \ldots, N_1\} = \emptyset$. Note that $J \cap I = \emptyset$ since $I \subseteq \{1, \ldots, N_1\}$. Therefore, we compute

$$\Big\|\sum_{j=1}^{N_2} v_{\phi(j)} - v_0\Big\| = \Big\|\sum_{j=1}^{N_1} v_j + \sum_{j \in J} v_j - v_0\Big\| \leq \Big\|\sum_{j=1}^{N_1} v_j - v_0\Big\| + \Big\|\sum_{j \in J} v_j\Big\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

giving convergence of $\sum_{j=1}^{\infty} v_{\phi(j)}$ to v_0. ■

As with series in $\mathbb{R}$, one of the essential features of absolutely convergent series is that their convergence is independent of rearrangement of terms. This mirrors the situation for series in $\mathbb{R}$.

6.4.5 Proposition (Absolute convergence implies unconditional Cauchy) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. For a sequence $(\mathrm{v_j})_{\mathrm{j} \in \mathbb{Z}_{>0}}$ consider the series $\mathrm{S} = \sum_{\mathrm{j}=1}^{\infty} \mathrm{v_j}$. If S is absolutely convergent then it is unconditionally Cauchy. Moreover, if S converges then, for any bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$, the series $\mathrm{S}_\phi = \sum_{\mathrm{j}=1}^{\infty} \mathrm{v}_{\phi(\mathrm{j})}$ converges absolutely to the same limit as S.*

Proof Let $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a bijection. First let us show that S_ϕ is absolutely convergent. Since S is absolutely convergent the sequence $(|S|_k)_{k \in \mathbb{Z}_{>0}}$ defined by

$$|S|_k = \sum_{j=1}^{k} \|v_j\|$$

is bounded and monotonically increasing. Thus there exists $M \in \mathbb{R}_{>0}$ such that $|S|_k \leq M$ for every $k \in \mathbb{Z}_{>0}$. Now define the sequence $(|S_\phi|_k)_{k \in \mathbb{Z}_{>0}}$ by

$$|S_\phi|_k = \sum_{j=1}^{k} v_{\phi(j)}.$$

For $k \in \mathbb{Z}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that $\{\phi(1), \ldots, \phi(k)\} \subseteq \{1, \ldots, N\}$. Then

$$|S_\phi|_k \leq \sum_{j=1}^{N} \|v_j\| \leq M.$$

Thus $(|S_\phi|_k)_{k \in \mathbb{Z}_{>0}}$ is bounded and monotonically increasing, and so convergent. Thus S_ϕ is absolutely convergent.

Next we show that if S is absolutely convergent then it is unconditionally Cauchy. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N_1 \in \mathbb{Z}_{>0}$ be such that

$$\sum_{j=N_1}^{\infty} \|v_j\| < \epsilon.$$

Now let $N_2 \in \mathbb{Z}_{>0}$ be such that $\{\phi(1), \dots, \phi(N_1)\} \subseteq \{1, \dots, N_2\}$. Let $k, l \geq N_2$ with $l > k$ and note that if $j \in \{k+1, \dots, l\}$ then $\phi^{-1}(j) \geq N_1$. Thus

$$\Big\| \sum_{j=k+1}^{l} v_{\phi(j)} \Big\| \leq \sum_{j=l+1}^{k} \|v_{\phi(j)}\| \leq \sum_{j=N_1}^{\infty} \|v_j\| < \epsilon,$$

showing that S_ϕ is Cauchy.

Now suppose that S converges to v_0 and let us show that S_ϕ converges to v_0. For $\epsilon \in \mathbb{R}_{>0}$ let $N_1 \in \mathbb{Z}_{>0}$ be such that

$$\Big\| \sum_{j=1}^{N_1} v_j - v_0 \Big\| < \frac{\epsilon}{2}$$

(this is possible since S converges to v_0) and such that

$$\sum_{j=N_1}^{\infty} \|v_j\| < \frac{\epsilon}{2} \tag{6.1}$$

(this is possible since S is absolutely convergent). There then exists $N_2 \in \mathbb{Z}_{>0}$ such that $\{\phi(1), \dots, \phi(N_1)\} \subseteq \{1, \dots, N_2\}$. Then

$$\sum_{j=1}^{N_2} v_{\phi(j)} = \sum_{j=1}^{N_1} v_j + \sum_{j \in J} v_{\phi(j)},$$

where $J = \{1, \dots, N_2\} \setminus \{\phi(1), \dots, \phi(N_1)\}$. Note that

$$\sum_{j \in J} \|v_{\phi(j)}\| \leq \sum_{j=N_1}^{\infty} \|v_j\| < \frac{\epsilon}{2}$$

by (6.1). Then

$$\begin{aligned}
\Big\| \sum_{j=1}^{N_2} v_{\phi(j)} - v_0 \Big\| &= \Big\| \sum_{j=1}^{N_1} v_j + \sum_{j \in J} v_{\phi(j)} - v_0 \Big\| \\
&\leq \Big\| \sum_{j=1}^{N_1} v_j - v_0 \Big\| + \Big\| \sum_{j \in J} v_{\phi(j)} \Big\| \\
&\leq \frac{\epsilon}{2} + \sum_{j \in J} \|v_{\phi(j)}\| \leq \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,
\end{aligned}$$

giving convergence of S_ϕ to v_0 as desired. ■

Thus Proposition 6.4.5 says that absolute convergence implies unconditional convergence. We shall see below in Theorem 6.4.8 that the two notions are equivalent if and only if the normed vector space is finite-dimensional. Thus the notion of unconditional convergence is the more general notion, and one may wonder whether absolute convergence is important. It is, and here is why.

6.4.6 Theorem (Absolute convergence and completeness) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. A normed $\mathbb{F}$-vector space $(\mathsf{V}, \|\cdot\|)$ is complete if and only if every absolutely convergent series in V converges.*

Proof Suppose that V is complete and let $\sum_{j=1}^{\infty} v_j$ be an absolutely convergent series. From Proposition 6.4.5 it follows that $\sum_{j=1}^{\infty} v_j$ is Cauchy, and so it converges since V is complete.

Now suppose that every absolutely convergent series converges, and let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence. Choose a subsequence $(v_{j_k})_{k\in\mathbb{Z}_{>0}}$ for which $\|u_{j_{k+1}} - u_{j_k}\| < \frac{1}{2^{k+1}}$. Then define $u_1 = v_{k_1}$ and $u_k = v_{j_k} - v_{j_{k-1}}$ so that the series $\sum_{k=1}^{\infty} u_k$ is absolutely convergent, and so convergent. This means therefore that

$$\lim_{k\to\infty} \|u_k\| = \lim_{k\to\infty} \|v_{j_k} - v_{j_{k-1}}\| = 0.$$

Thus the sequence $(v_{j_k})_{k\in\mathbb{Z}_{>0}}$ is convergent. Suppose it converges to v. Now, for $\epsilon > 0$ choose k and j sufficiently large that $\|v_j - v_{j_k}\| < \frac{\epsilon}{2}$ and $\|v - v_{j_k}\| < \frac{\epsilon}{2}$. Then we have

$$\|v - v_j\| \leq \|v - v_{j_k}\| + \|v_{j_k} - v_j\| < \epsilon,$$

so showing that $(v_j)_{j\in\mathbb{Z}_{>0}}$ converges to v. ■

The following trivial corollary is sometimes useful by itself.

6.4.7 Corollary (Absolutely convergent sequences in Banach spaces converge) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. If $(\mathsf{V}, \|\cdot\|)$ is a $\mathbb{F}$-Banach space and if $\sum_{\mathrm{j}=1}^{\infty} \mathrm{v}_{\mathrm{j}}$ is an absolutely convergent series in V, then $\sum_{\mathrm{j}=1}^{\infty} \mathrm{v}_{\mathrm{j}}$ is convergent.*

Now let us explore the possibility of a converse to Proposition 6.4.5. That is, let us consider the question, “Is it true that an unconditionally convergent series is absolutely convergent?” In Theorem 2.4.5 we saw that this was true for series in $\mathbb{R}$. However, this is not generally true in normed vector spaces, but holds if and only if the vector space is finite-dimensional. This is an instance of where the difference between finite- and infinite-dimensions shows up.

6.4.8 Theorem (Absolute convergence and unconditional Cauchy agree (only) in finite-dimensions) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. Then the set of absolutely convergent series and the set of unconditionally Cauchy series coincide if and only if V is finite-dimensional.*

Proof From Proposition 6.4.5 we know that absolutely convergent series are always unconditionally convergent. Suppose that V is finite-dimensional and that $\sum_{j=1}^{\infty} v_j$ is unconditionally convergent. Let us also suppose that $\mathbb{F} = \mathbb{R}$ for the moment. Choose a basis $\{e_1, \dots, e_n\}$ for V and write

$$v_j = v_j^1 e_1 + \cdots + v_j^n e_n$$

for $v_j^l \in \mathbb{F}$, $j \in \mathbb{Z}_{>0}$, $l \in \{1, \dots, n\}$. By Theorem 6.1.15 we can use any norm on V we wish to discuss convergence, so let us use the ∞-norm induced by the basis:

$$\|v^1 e_1 + \cdots + v^n e_n\| = \max\{|v^1|, \dots, |v^n|\}.$$

Let $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a bijection so that $\sum_{j=1}^{\infty} v_{\phi(j)}$ converges, say to $v_0 \in \mathsf{V}$. Let us write

$$v_0 = v_0^1 e_1 + \cdots + v_0^n e_n.$$

Now let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that

$$\Big\| \sum_{j=1}^{N} v_{\phi(j)} - v_0 \Big\| < \epsilon.$$

Then

$$\Big| \sum_{j=1}^{N} v_{\phi(j)}^l - v_0^l \Big| \le \Big\| \sum_{j=1}^{N} v_{\phi(j)} - v_0 \Big\| < \epsilon.$$

Thus $\sum_{j=1}^{\infty} v_{\phi(j)}^l$ converges to v_0^l for each $l \in \{1,\dots,n\}$. Thus $\sum_{j=1}^{\infty} v_j^l$ is unconditionally convergent, and so absolutely convergent by Theorem 2.4.5. Now again let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that

$$\sum_{j=N+1}^{\infty} |v_j^l| < \epsilon, \qquad l \in \{1,\dots,n\},$$

this being possible by absolute convergence of $\sum_{j=1}^{\infty} v_j^l$. Then, for any $l \in \{1,\dots,n\}$,

$$\sum_{j=N+1}^{\infty} \|v_j\| \le \sum_{j=N+1}^{\infty} |v_j^l| < \epsilon,$$

giving absolute convergence of $\sum_{j=1}^{\infty} v_j$.

If V is a finite-dimensional $\mathbb{C}$-vector space, then it is also a finite-dimensional $\mathbb{R}$-vector space of twice the dimension, and so the above arguments can be used to show that an unconditionally convergent sum is absolutely convergent.

It remains to show that if V is infinite-dimensional then there exists an unconditionally convergent series that is not absolutely convergent. We do this via a sequence of lemmata, the first of which seems to have nothing to do with the problem at hand. Let us suppose that $\mathbb{F} = \mathbb{R}$.

The following lemma is crucial, and is called the ***Dvoretzky–Rogers Lemma***.

1 Lemma *Let* $\mathrm{C} \subseteq \mathbb{R}^{\mathrm{n}}$ *be a compact convex set with nonempty interior and with centre at* $\mathbf{0}_{\mathbb{R}^{\mathrm{n}}}$ *and let* $\mathrm{k} \in \{1,\dots,\mathrm{n}\}$*. Then there exists* $\mathbf{x}_1,\dots,\mathbf{x}_{\mathrm{n}} \in \mathrm{bd}(\mathrm{C})$ *such that, for any* $\lambda_1,\dots,\lambda_{\mathrm{k}} \in \mathbb{R}$*,*

$$\lambda_1 \mathbf{x}_1 + \cdots + \lambda_{\mathrm{k}} \mathbf{x}_{\mathrm{k}} \in \lambda \mathrm{C} \triangleq \{\lambda \mathbf{x} \mid \mathbf{x} \in \mathrm{C}\},$$

where

$$\lambda^2 = \Big(2 + \tfrac{\mathrm{k}(\mathrm{k}-1)}{\mathrm{n}}\Big)(\lambda_1^2 + \cdots + \lambda_{\mathrm{k}}^2).$$

Proof By Theorem **??** let E be the ellipsoid with largest volume contained in C. If $\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ is invertible then hypotheses of the lemma hold for the convex set $\boldsymbol{A}(C)$ and the conclusions hold for the points $\boldsymbol{A}\boldsymbol{x}_1,\dots,\boldsymbol{A}\boldsymbol{x}_n$. Thus we can apply an invertible linear transformation of $\mathbb{R}^n$ to the problem without changing either the hypotheses or the conclusions. Let us suppose that $\boldsymbol{A}$ has been chosen such that $\boldsymbol{A}(E) = \overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n})$, the closed unit ball in the 2-norm in $\mathbb{R}^n$. For the remainder of the proof we work with the transformed problem.

We next claim that there exists an orthogonal matrix $\boldsymbol{R}$, thought of as a linear mapping from $\mathbb{R}^n$ to itself, and points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in \overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n}) \cap C$ such that

$$\boldsymbol{y}_j \triangleq \boldsymbol{R}\boldsymbol{x}_j = (y_j^1, \dots, y_j^j, 0, \dots, 0), \qquad j \in \{1, \dots, n\}, \tag{6.2}$$

(i.e., the last $n - j$ components of $\boldsymbol{R}\boldsymbol{x}_j$ are zero) and such that

$$(y_j^1)^2 + \dots + (y_j^{j-1})^2 = 1 - (y_j^j)^2 \le \tfrac{j-1}{n}, \qquad j \in \{1, \dots, n\}. \tag{6.3}$$

We construct the points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ inductively. For $j = 1$ we take $\boldsymbol{x}_1 \in \overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n}) \cap C$ (this is possible by our initial definition of E). We then make an orthogonal change of basis for which $\boldsymbol{x}_1$ is the first basis vector. This defines an orthogonal transformation $\boldsymbol{R}_1$ satisfying (6.2) and (6.3) for $j = 1$. Suppose now, for $k - 1 < n$, that we have defined $\boldsymbol{R}_{k-1}$ and $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-1} \in \overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n}) \cap C$ satisfying (6.2) and (6.3) for $j \in \{1, \dots, k-1\}$. Define $f_k \colon \mathbb{R}_{\ge 0} \times \mathbb{R}^n \to \mathbb{R}$ by

$$f_k(\epsilon, \boldsymbol{x}) = (1 + \epsilon)^{n-k+1}((y^1)^2 + \dots + (y^{k-1})^2) + (1 + \epsilon + \epsilon^2)^{-k+1}((y^k)^2 + \dots + (y^n)^2),$$

where $\boldsymbol{y} = \boldsymbol{R}_{k-1}\boldsymbol{x}$. For $\epsilon \in \mathbb{R}_{\ge 0}$ define

$$E_\epsilon = \{\boldsymbol{x} \in \mathbb{R}^n \mid f(\epsilon, \boldsymbol{x}) \le 1\}.$$

Thus E_ϵ is an ellipsoid. We claim that for $\epsilon \in \mathbb{R}_{>0}$ the volume of E_ϵ exceeds that of $\overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n})$. To see this, consider the linear transformation $\boldsymbol{T}_\epsilon$ of $\mathbb{R}^n$ defined by

$$\begin{aligned}\boldsymbol{T}_\epsilon(y^1, \dots, y^n) = (&\sqrt{(1 + \epsilon)^{n-k+1}}y^1, \dots, \sqrt{(1 + \epsilon)^{n-k+1}}y^{k-1}, \\ &\sqrt{(1 + \epsilon + \epsilon^2)^{-k+1}}y^k, \dots, \sqrt{(1 + \epsilon + \epsilon^2)^{-k+1}}y^n).\end{aligned}$$

Thus $\boldsymbol{T}_\epsilon(E_\epsilon) = \overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n})$. Using the change of variables formula for the integral in $\mathbb{R}^n$ we have the volume of E_ϵ as $\det \boldsymbol{T}_\epsilon^{-1}$ times the volume of $\mathsf{B}(1, \mathbf{0}_{\mathbb{R}^n})$. Since

$$\det \boldsymbol{T}_\epsilon^{-1} = \Big(\frac{1 + \epsilon + \epsilon^2}{1 + \epsilon}\Big)^{(n-k+1)(k-1)/2} > 1$$

for $\epsilon \in \mathbb{R}_{>0}$, we indeed have the volume of E_ϵ as exceeding that of $\overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n})$.

Now, since $\mathsf{B}(1, \mathbf{0}_{\mathbb{R}^n})$ is the largest ellipsoid contained in C, there exists a point $\boldsymbol{x}_\epsilon \in \operatorname{bd}(C) \cap E_\epsilon$. Since $\boldsymbol{x}_\epsilon \in \operatorname{bd}(C)$ and since $\overline{\mathsf{B}}(1, \mathbf{0}_{\mathbb{R}^n}) \subseteq C$ it follows that $\|\boldsymbol{x}_\epsilon\| \ge 1$ (where $\|\cdot\|$ is the 2-norm on $\mathbb{R}^n$). Letting $\boldsymbol{y}_\epsilon = \boldsymbol{R}_{k-1}\boldsymbol{x}_\epsilon$ we have

$$((y_\epsilon^1)^2 + \dots + (y_\epsilon^{k-1})^2) + ((y_\epsilon^k)^2 + \dots + (y_\epsilon^n)^2) \ge 1.$$

Subtracting this inequality from the inequality $f(\epsilon, \boldsymbol{x}_\epsilon) \le 1$ gives

$$\begin{aligned}((1 + \epsilon)^{n-k+1} - 1)&((y_\epsilon^1)^2 + \dots + (y_\epsilon^{k-1})^2) \\ &+ ((1 + \epsilon + \epsilon^2)^{-k+1} - 1)((y_\epsilon^k)^2 + \dots + (y_\epsilon^n)^2) \le 0.\end{aligned} \tag{6.4}$$

Let $(\epsilon_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathbb{R}_{>0}$ converging to zero. The resulting sequence $(\boldsymbol{x}_{\epsilon_j})_{j\in\mathbb{Z}_{>0}}$ is in $\mathrm{bd}(C)$ which is compact, being a closed subset of a compact set (Corollary **??**). Therefore, by the Bolzano–Weierstrass Theorem, there exists a subsequence of $(\boldsymbol{x}_{\epsilon_j})_{j\in\mathbb{Z}_{>0}}$ converging to some $\boldsymbol{x}_0 \in \overline{\mathsf{B}}(1, \boldsymbol{0}_{\mathbb{R}^n}) \cap \mathrm{bd}(C)$. Moreover, denoting $\boldsymbol{y}_0 = \boldsymbol{R}\boldsymbol{x}_0$, (6.4) gives

$$\begin{aligned}
&\frac{1}{\epsilon}((1+\epsilon)^{n-k+1}-1)((y_\epsilon^1)^2+\cdots+(y_\epsilon^{k-1})^2)\\
&\quad+((1+\epsilon+\epsilon^2)^{-k+1}-1)((y_\epsilon^k)^2+\cdots+(y_\epsilon^n)^2)\le 0\\
\implies\quad &\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}((1+\epsilon)^{n-k+1}-1)((y_\epsilon^1)^2+\cdots+(y_\epsilon^{k-1})^2)\\
&\quad+((1+\epsilon+\epsilon^2)^{-k+1}-1)((y_\epsilon^k)^2+\cdots+(y_\epsilon^n)^2)\le 0\\
\implies\quad &(n-k+1)((y_0^1)^2+\cdots+(y_0^{k-1})^2)+(-k+1)((y_0^k)^2+\cdots+(y_0^n)^2)\le 0.
\end{aligned}$$

Now define $\boldsymbol{x}_k = \boldsymbol{x}_0$. If $\boldsymbol{R}_{k-1}\boldsymbol{x}_k \in \mathrm{span}_{\mathbb{R}}(\boldsymbol{e}_1,\dots,\boldsymbol{e}_{k-1})$ then clearly the last $n-k$ components of $\boldsymbol{R}_{k-1}\boldsymbol{x}_k$ are zero in the basis defined by $\boldsymbol{R}_{k-1}$. If not, then the vectors $\{\boldsymbol{R}_{k-1}^{-1}\boldsymbol{e}_1,\dots,\boldsymbol{R}_{k-1}^{-1}\boldsymbol{e}_{k-1},\boldsymbol{x}_k\}$ span a subspace of dimension k and by choosing an orthogonal complement in this subspace to $\mathrm{span}_{\mathbb{R}}(\boldsymbol{R}_{k-1}^{-1}\boldsymbol{e}_1,\dots,\boldsymbol{R}_{k-1}^{-1}\boldsymbol{e}_{k-1})$ we get an orthonormal basis for $\mathbb{R}^n$ where the first $k-1$ basis vectors are those defined by $\boldsymbol{R}_{k-1}$ and the first k basis vectors span a subspace containing $\boldsymbol{x}_k$. Thus the last $n-k$ components of $\boldsymbol{x}_k$ in this basis will be zero, and the components of $\boldsymbol{x}_1,\dots,\boldsymbol{x}_{k-1}$ will be unchanged from those in the basis defined by $\boldsymbol{R}_{k-1}$. This new orthonormal basis defines an orthogonal matrix $\boldsymbol{R}_k$. This gives condition (6.2). Moreover, if we abuse notation slightly and denote by $(y^1,\dots,y^n)$ the coordinates in the basis defined by $\boldsymbol{R}_k$, the point $\boldsymbol{y}_k = \boldsymbol{R}_k\boldsymbol{x}_k$ satisfies

$$(n-k+1)((y_k^1)^2+\cdots+(y_k^{k-1})^2)+(-k+1)(y_k^k)^2\le 0.$$

Since we also have $(y_k^1)^2+\cdots+(y_k^k)^2 = 1$ we then get

$$(y_k^1)^2+\cdots+(y_k^{k-1})^2 = 1-(y_k^k)^2 = \frac{k-1}{n},$$

and so (6.3) also holds.

Finally, let $\lambda_1,\dots,\lambda_k\in\mathbb{R}$. We compute the square of the length of $\lambda_1\boldsymbol{x}_1+\cdots+\lambda_k\boldsymbol{x}_k$ as

$$\begin{aligned}
\sum_{j=1}^k\Big(\sum_{l=1}^k\lambda_l y_l^j\Big)^2 &\le \sum_{j=1}^k\Big(2\lambda_j^2(y_j^j)^2+2\Big(\sum_{l=j+1}^k\lambda_l y_l^j\Big)^2\Big)\\
&\le 2\sum_{j=1}^k\Big(\lambda_j^2(y_j^j)^2+\Big(\sum_{l=j+1}^k\lambda_l^2\Big)\Big(\sum_{m=j+1}^k(y_m^j)^2\Big)\Big)\\
&= 2\sum_{j=1}^k\Big((y_j^j)^2+\sum_{l=1}^k\sum_{m=1}^{\min\{j-1,l-1\}}(y_l^m)^2\Big)\lambda_j^2.
\end{aligned}$$

Since (6.3) holds we have

$$(y_j^j)^2+\sum_{l=1}^k\sum_{m=1}^{\min\{j-1,l-1\}}(y_l^m)^2\le 1+\sum_{l=1}^k\frac{l-1}{n},\qquad j\in\{1,\dots,k\}.$$

Therefore, the length of $\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k$ is bounded above by

$$\sum_{j=1}^{k} \Big(1 + \sum_{l=1}^{k} \frac{l-1}{n}\Big)\lambda_j^2 = \Big(2 + \frac{k(k-1)}{n}\Big) \sum_{j=1}^{k} \lambda_j^2.$$

In other words, $\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k \in \overline{\mathsf{B}}(\lambda, \boldsymbol{0}_{\mathbb{R}^n})$ where

$$\lambda^2 = \Big(2 + \frac{k(k-1)}{n}\Big) \sum_{j=1}^{k} \lambda_j^2.$$

Thus $\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k \in \lambda C$ since $\overline{\mathsf{B}}(1, \boldsymbol{0}_{\mathbb{R}^n}) \subseteq C$. ▼

2 Lemma *Let* $(\mathsf{V}, \|\cdot\|)$ *be an infinite-dimensional normed* $\mathbb{R}$*-vector space, let* $\mathrm{k} \in \mathbb{Z}_{>0}$*, and let* $\mathrm{c}_1, \dots, \mathrm{c}_\mathrm{k} \in \mathbb{R}_{>0}$*. Then there exists* $\mathrm{v}_1, \dots, \mathrm{v}_\mathrm{k} \in \mathsf{V}$ *such that*

(i) $\|\mathrm{v}_\mathrm{j}\|^2 = \mathrm{c}_\mathrm{j}$, $\mathrm{j} \in \{1, \dots, \mathrm{k}\}$*, and*

(ii) $\big\|\sum_{\mathrm{j}\in\mathrm{J}} \mathrm{v}_\mathrm{j}\big\|^2 \leq 3 \sum_{\mathrm{j}\in\mathrm{J}} \mathrm{c}_\mathrm{j}$ *for every subset* $\mathrm{J} \subseteq \{1, \dots, \mathrm{k}\}$*.*

Proof Let $n = k(k-1)$ and let $u_1, \dots, u_n \in \mathsf{V}$ be linearly independent. Define

$$C = \Big\{(x^1, \dots, x^n) \in \mathbb{R}^n \;\Big|\; \|x^1 u_1 + \cdots + x^n u_n\| \leq 1\Big\}.$$

We claim that C is convex, compact, has nonempty interior, and has centre $\boldsymbol{0}_{\mathbb{R}^n}$. One sees this as follows. The map

$$\mathsf{L}\colon (x^1, \dots, x^n) \mapsto x^1 u_1 + \cdots + x^n u_n$$

is a linear injection of $\mathbb{R}^n$ onto the n-dimensional subspace spanned by $u_1, \dots, u_n$. One can then define a norm on $\mathbb{R}^n$ to be the norm induced from the restriction of the norm in V to the subspace $\mathsf{L}(\mathbb{R}^n)$. The closed unit ball in this norm is simply C. Then $\mathsf{L}(C)$ is the intersection of the closed unit ball in V with the subspace $\mathsf{L}(\mathbb{R}^n)$. Thus $\mathsf{L}(C)$ is the intersection of convex sets and so is convex by Exercise **??**. Moreover, $\mathsf{L}(C)$ is clearly a closed and bounded subset of $\mathsf{L}(\mathbb{R}^n)$ and so is compact by the Heine–Borel Theorem. The unit ball in any norm clearly has nonempty interior (see Exercise 6.1.1). Also, $\boldsymbol{0}_{\mathbb{R}^n}$ is the centre of C since $\boldsymbol{x} \in C$ if and only if $-\boldsymbol{x} \in C$.

Let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ be as in Lemma 1 and define

$$v_j = \sqrt{c_j}\mathsf{L}(\boldsymbol{x}_j), \qquad j \in \{1, \dots, k\},$$

where $\mathsf{L}\colon \mathbb{R}^n \to \mathsf{V}$ is the map from the preceding paragraph. Then

$$\|v_j\|^2 = c_j \|x_j^1 u_1 + \cdots + x_j^n u_n\| = c_j, \qquad j \in \{1, \dots, k\},$$

since $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in \mathrm{bd}(C)$. Now let $J \subseteq \{1, \dots, k\}$. Then, by Lemma 1,

$$\sum_{j\in J} \sqrt{c_j}\boldsymbol{x}_j \in \lambda C$$

where

$$\lambda^2 = \Big(2 + \frac{k(k-1)}{n}\Big) \sum_{j\in J} c_j = 3 \sum_{j\in J} c_j.$$

This implies that

$$\mathsf{L}\Big(\sum_{j\in J}\sqrt{c_j}\boldsymbol{x}_j\Big)\in\mathsf{L}(\lambda C)\quad\Longrightarrow\quad\Big\|\sum_{j\in J}v_j\Big\|\leq\Big(3\sum_{j\in J}c_j\Big)^{1/2},$$

as claimed. ▼

3 Lemma *Let* $(\mathsf{V},\|\cdot\|)$ *be an infinite-dimensional normed* $\mathbb{R}$*-vector space and let* $\sum_{j=1}^{\infty}c_j$ *be a convergent series in* $\mathbb{R}_{>0}$*. Then there exists an unconditionally Cauchy series* $\sum_{j=1}^{\infty}v_j$ *in* V *such that* $\|v_j\|^2=c_j$, $j\in\mathbb{Z}_{>0}$*.*

Proof Define $n_0=0$ and define n_1 such that

$$\Big(\sum_{j=n_1+1}^{\infty}c_j\Big)^{1/2}<1,$$

this being possible since $\sum_{j=1}^{\infty}c_j$ is a convergent series of positive terms. Then define $n_2>n_1$ such that

$$\Big(\sum_{j=n_2+1}^{\infty}c_j\Big)^{1/2}<\frac{1}{4}.$$

Carrying on in this way we define an increasing sequence $(n_j)_{j\in\mathbb{Z}_{\geq0}}$ such that

$$\Big(\sum_{j=n_k+1}^{n_{k+1}}c_j\Big)^{1/2}<\Big(\sum_{j=n_k+1}^{\infty}c_j\Big)^{1/2}<\frac{1}{k^2},\qquad k\in\mathbb{Z}_{>0}.$$

The series

$$\sum_{k=0}^{\infty}\Big(\sum_{j=n_k+1}^{n_{k+1}}c_j\Big)^{1/2}$$

then converges by Example 2.4.2–**??**. Take $k\in\mathbb{Z}_{\geq0}$. By Lemma 2 let v_j, $j\in\{n_k+1,\dots,n_{k+1}\}$, be such that $\|v_j\|^2=c_j$ and such that

$$\Big\|\sum_{j\in J}v_j\Big\|^2\leq3\sum_{j\in J}c_j$$

for any $J\subseteq\{n_k+1,\dots,n_{k+1}\}$. Let $\epsilon\in\mathbb{R}_{>0}$ and choose $N_1\in\mathbb{Z}_{>0}$ such that

$$\sum_{k=N_1}^{\infty}\Big(\sum_{j=n_k+1}^{n_{k+1}}c_j\Big)^{1/2}<\frac{\epsilon}{3}.$$

Let $\phi\colon\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}$ be a bijection and choose $N_2\in\mathbb{Z}_{>0}$ such that

$$\{1,\dots,n_{N_1}\}\subseteq\{\phi(1),\dots,\phi(N_2)\}.$$

Thus

$$(v_{\phi(j)})_{j=N_2}^{\infty}\subseteq(v_j)_{j=N_1+1}^{\infty}.$$

Let $N_3 > N_2$ and let $k \geq N_1$. Denote by $J_k \subseteq \{n_k + 1, \dots, n_{k+1}\}$ the indices such that $j \in J_k$ if and only if $\phi(j) \in \{N_2, \dots, N_3\}$. Then we have

$$\Big\|\sum_{j=N_2}^{N_3} v_{\phi(j)}\Big\| = \Big\|\sum_{k=N_1}^{\infty}\sum_{j\in J_k} v_j\Big\| \leq \sum_{k=N_1}^{\infty}\Big\|\sum_{j\in J_k} v_j\Big\| \leq \sum_{k=N_1}^{\infty}\Big(3\sum_{j=n_k+1}^{n_k} c_j\Big)^{1/2} < \epsilon.$$

Thus the norm of the N_3rd partial sum minus the N_2nd partial sum for the series $\sum_{j=1}^{\infty} v_{\phi(j)}$ is less than ϵ. Thus this series is Cauchy and so $\sum_{j=1}^{\infty} v_j$ is unconditionally Cauchy. ▼

Now let us prove the theorem. Consider the sequence $\big(c_j = \frac{1}{j^2}\big)_{j\in\mathbb{Z}_{>0}}$ and by Lemma 3 let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence for which $\|v_j\|^2 = c_j$ and for which the series $\sum_{j=1}^{\infty} v_j$ is unconditionally Cauchy. But $\sum_{j=1}^{\infty}\|v_j\| = \sum_{j=1}^{\infty}\frac{1}{j}$ is divergent by Example 2.4.2–?? and so $\sum_{j=1}^{\infty} v_j$ is not absolutely convergent. This proves the theorem for normed $\mathbb{R}$-vector spaces. For normed $\mathbb{C}$-vector spaces we note that these are also normed $\mathbb{R}$-vector spaces. Since none of the constructions in the proof alter when complex scalars are replaced with real scalars, the proof is also valid for normed $\mathbb{C}$-vector spaces. ■

6.4.3 Algebraic operations on series

Let us close by indicating that convergence of series respects the algebraic structure of vector spaces. We first give two definitions of products of series of scalars and vectors.

6.4.9 Definition (Scalar multiplication of series) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Let $S = \sum_{j=0}^{\infty} v_j$ be a series in V and let $s = \sum_{j=0}^{\infty} a_j$ be series in $\mathbb{R}$.

(i) The ***product*** of s and S is the double series $\sum_{j,k=0}^{\infty} a_j v_k$.

(ii) The ***Cauchy product*** of s and S is the series $\sum_{k=0}^{\infty}\Big(\sum_{j=0}^{k} a_j v_{k-j}\Big)$. •

Now we can state the interaction between convergence of series and the vector space operations.

6.4.10 Proposition (Algebraic operations on series) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Let $\mathrm{S} = \sum_{\mathrm{j}=0}^{\infty} \mathrm{u_j}$ and $\mathrm{T} = \sum_{\mathrm{j}=0}^{\infty} \mathrm{v_j}$ be series in V converging to $\mathrm{U_0}$ and $\mathrm{V_0}$, respectively, let $\mathrm{s} = \sum_{\mathrm{j}=0}^{\infty} \mathrm{a_j}$ be a series in $\mathbb{F}$ converging to $\mathrm{A_0}$, and let $\mathrm{a} \in \mathbb{F}$. Then the following statements hold:*

(i) the series $\sum_{\mathrm{j}=0}^{\infty} \mathrm{av_j}$ converges to $\mathrm{aV_0}$;

(ii) the series $\sum_{\mathrm{j}=0}^{\infty}(\mathrm{u_j} + \mathrm{v_j})$ converges to $\mathrm{U_0} + \mathrm{V_0}$;

(iii) if s and T are absolutely convergent, then the product of s and T is absolutely convergent and converges to $\mathrm{A_0V_0}$;

(iv) if s and T are absolutely convergent, then the Cauchy product of s and T is absolutely convergent and converges to $\mathrm{A_0V_0}$;

(v) if s or T are absolutely convergent, then the Cauchy product of s and T is convergent and converges to $\mathrm{A_0V_0}$.

Proof (i) Since $\sum_{j=0}^{k} av_j = a\sum_{j=0}^{k} v_j$, this follows from part (i) of Proposition 6.2.6.

(ii) Since $\sum_{j=0}^{\infty}(u_j + v_j) = \sum_{j=0}^{k} u_j + \sum_{j=0}^{k} v_j$, this follows from part (ii) of Proposition 6.2.6.

(iii) and (iv) To prove these parts of the result, we first make a general argument. We note that $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ is a countable set (e.g., by Proposition **??**), and so there exists a bijection, in fact many bijections, $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$. For such a bijection ϕ, suppose that we are given a double sequence $(v_{jk})_{j,k\in\mathbb{Z}_{\geq 0}}$ and define a sequence $(v_j^{\phi})_{j\in\mathbb{Z}_{>0}}$ by $v_j^{\phi} = x_{kl}$ where $(k,l) = \phi(j)$. We then claim that, for any bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0}$, the double series $A = \sum_{k,l=1}^{\infty} v_{kl}$ converges absolutely if and only if the series $A^{\phi} = \sum_{j=1}^{\infty} v_j^{\phi}$ converges absolutely.

Indeed, suppose that the double series $\|A\| = \sum_{k,l=1}^{\infty}\|v_{kl}\|$ converges to $\beta \in \mathbb{R}$. For $\epsilon > 0$ the set

$$\{(k,l) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \mid |\|A\|_{kl} - \beta| \geq \epsilon\}$$

is then finite. Therefore, there exists $N \in \mathbb{Z}_{>0}$ such that, if $(k,l) = \phi(j)$ for $j \geq N$, then $|\|A\|_{kl} - \beta| < \epsilon$. It therefore follows that $|\|A^{\phi}\|_j - \beta| < \epsilon$ for $j \geq N$, where $\|A^{\phi}\|$ denotes the series $\sum_{j=1}^{\infty}|v_j^{\phi}|$. This shows that the series $\|A^{\phi}\|$ converges to β.

For the converse, suppose that the series $\|A^{\phi}\|$ converges to β. Then, for $\epsilon > 0$ the set

$$\{j \in \mathbb{Z}_{>0} \mid \|A^{\phi}|_j - \beta| \geq \epsilon\}$$

is finite. Therefore, there exists $N \in \mathbb{Z}_{>0}$ such that

$$\{(k,l) \in \mathbb{Z}_{\geq 0} \mid k,l \geq N\} \cap \{(k,l) \in \mathbb{Z}_{\geq 0} \mid \|A^{\phi}|_{\phi^{-1}(k,l)} - \beta| \geq \epsilon\} = \emptyset.$$

It then follows that for $k,l \geq N$ we have $\|A|_{kl} - \beta| < \epsilon$, showing that $|A|$ converges to β.

Thus we have shown that A is absolutely convergent if and only if A^{ϕ} is absolutely convergent for any bijection $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$. From Proposition 6.4.5 we know that the limit of an absolutely convergent series or double series is independent of the manner in which the terms in the series are arranged.

Consider now a term in the product of s and T. It is easy to see that this term appears exactly once in the Cauchy product of s and T. Conversely, each term in the Cauchy product appears exactly one in the product. Thus the product and Cauchy product are simply rearrangements of one another. Moreover, each term in the product and the Cauchy product appears exactly once in the expression

$$\Big(\sum_{j=0}^{N} a_j\Big)\Big(\sum_{k=0}^{N} v_k\Big)$$

as we allow N to go to ∞. That is to say,

$$\sum_{j,k=0}^{\infty} a_j v_k = \sum_{k=0}^{\infty}\Big(\sum_{j=k}^{k} a_j v_{k-j}\Big) = \lim_{N\to\infty}\Big(\sum_{j=0}^{N} a_j\Big)\Big(\sum_{k=0}^{N} v_k\Big).$$

However, this last limit is exactly $A_0 V_0$, using part (iii) of Proposition 6.2.6.

(v) Suppose that s converges absolutely. Let $(s_k)_{k\in\mathbb{Z}_{>0}}$, $(T_k)_{k\in\mathbb{Z}_{>0}}$, and $((sT)_k)_{k\in\mathbb{Z}_{>0}}$ be the sequences of partial sums for s, T, and the Cauchy product, respectively. Also define $\tau_k = T_k - V_0$, $k \in \mathbb{Z}_{\geq 0}$. Then

$$\begin{aligned}(sT)_k &= a_0v_0 + (a_0v_1 + a_1v_0) + \cdots + (a_0v_k + \cdots + a_kv_0)\\ &= a_0T_k + a_1T_{k-1} + \cdots + a_kT_0\\ &= a_0(V_0 + \tau_k) + a_1(V_0 + \tau_{k-1}) + \cdots + a_k(V_0 + \tau_0)\\ &= s_kV_0 + a_0\tau_k + a_1\tau_{k-1} + \cdots + a_k\tau_0.\end{aligned}$$

Since $\lim_{k\to\infty} s_kV_0 = A_0V_0$ by part Proposition 2.4.30(**??**), this part of the result will follow if we can show that

$$\lim_{k\to\infty}(a_0\tau_k + a_1\tau_{k-1} + \cdots + a_k\tau_0) = 0. \tag{6.5}$$

Denote

$$\sigma = \sum_{j=0}^{\infty}|a_j|,$$

and for $\epsilon > 0$ choose $N_1 \in \mathbb{Z}_{>0}$ such that $\|\tau_j\| \leq \frac{\epsilon}{2\sigma}$ for $j \geq N_1$, this being possible since $(\tau_j)_{j\in\mathbb{Z}_{>0}}$ clearly converges to zero. Then, for $k \geq N_1$,

$$\begin{aligned}\|a_0\tau_k + a_1\tau_{k-1} + \cdots + a_k\tau_0\| &\leq \|a_0\tau_k + \cdots + a_{k-N_1-1}\tau_{N_1-1}\| + \|a_{k-N_1}\tau_{N_1} + \cdots + a_k\tau_0\|\\ &\leq \tfrac{\epsilon}{2} + \|a_{k-N_1}\tau_{N_1} + \cdots + a_k\tau_0\|.\end{aligned}$$

Since $\lim_{k\to\infty} a_k = 0$, choose $N_2 \in \mathbb{Z}_{>0}$ such that

$$\|a_{k-N_1}\tau_{N_1} + \cdots + a_k\tau_0\| < \tfrac{\epsilon}{2}$$

for $k \geq N_2$. Then

$$\begin{aligned}&\limsup_{k\to\infty}\|a_0\tau_k + a_1\tau_{k-1} + \cdots + a_k\tau_0\|\\ &\qquad= \lim_{k\to\infty}\sup\{\|a_0\tau_j + a_1\tau_{j-1} + \cdots + a_j\tau_0\| \mid j \geq k\}\\ &\qquad\leq \lim_{k\to\infty}\sup\{\tfrac{\epsilon}{2} + \|a_{k-N_1}\tau_{N_1} + \cdots + a_k\tau_0\| \mid j \geq k\}\\ &\qquad\leq \sup\{\tfrac{\epsilon}{2} + \|a_{k-N_1}\tau_{N_1} + \cdots + a_k\tau_0\| \mid j \geq N_2\} \leq \epsilon.\end{aligned}$$

Thus

$$\limsup_{k\to\infty}\|a_0\tau_k + a_1\tau_{k-1} + \cdots + a_k\tau_0\| \leq 0,$$

and since clearly

$$\liminf_{k\to\infty}\|a_0\tau_k + a_1\tau_{k-1} + \cdots + a_k\tau_0\| \geq 0,$$

we infer that (6.5) holds by Proposition 2.3.17.

If T converges absolutely, the above argument can be modified by defining

$$\sigma = \sum_{j=0}^{\infty}\|v_j\|$$

and swapping the rôles of s and T in the remainder of the proof. ■

6.4.4 Multiple series

One also has the notion of double series in normed vector spaces.

6.4.11 Definition (Double series) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be a $\mathbb{F}$-vector space. A ***double series*** in V is a sum of the form $\sum_{j,k=1}^{\infty} v_{jk}$ where $(v_{jk})_{j,k\in\mathbb{Z}_{>0}}$ is a double sequence in V. •

We then have the following notions of convergence of double series.

6.4.12 Definition (Convergence and absolute convergence of double series) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a $\mathbb{F}$-vector (semi)normed space. Let $(v_{jk})_{j,k\in\mathbb{Z}_{>0}}$ be a double sequence in V and consider the double series

$$S = \sum_{j,k=1}^{\infty} v_{jk}.$$

The corresponding sequence of ***partial sums*** is the double sequence $(S_{jk})_{j,k\in\mathbb{Z}_{>0}}$ defined by

$$S_{jk} = \sum_{l=1}^{j} \sum_{m=1}^{k} v_{lm}.$$

Let $v_0 \in \mathsf{V}$. The double series:

(i) ***converges to*** $\mathbf{v_0}$, and we write $\sum_{j,k=1}^{\infty} v_{jk} = v_0$, if the double sequence of partial sums converges to v_0;

(ii) has v_0 as a ***limit*** if it converges to v_0;

(iii) is ***convergent*** if it converges to some member of V;

(iv) ***converges absolutely***, or is ***absolutely convergent***, if the series

$$\sum_{j,k=1}^{\infty} \|v_{jk}\|$$

converges;

(v) ***converges conditionally***, or is ***conditionally convergent***, if it is convergent, but not absolutely convergent;

(vi) ***diverges*** if it does not converge. •

6.4.5 Cesàro convergence of sequences and series

If a sequence diverges, all hope may not be lost. Indeed, it is possible that convergence may not actually be what one was interested in. This seems a somewhat absurd proposition at first glance, but it actually forms the first steps towards a powerful theory of Fourier series, as we shall see in Section 12.2.7. The point is that when one has a divergent sequence or series, one should not just throw in the towel. It is possible that by modifying one's notion of convergence, useful information can still be extracted.

The idea of Cesàro convergence is that one should average the sequence and see if the averaged sequence converges. The same idea can be applied to sums via their partial sums.

6.4.13 Definition (Cesàro[1] convergence) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space, and let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in V.

(i) The ***Cesàro means*** for the sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ is the sequence $(\bar{v}^1_k))_{k\in\mathbb{Z}_{>0}}$ where

$$\bar{v}^1_k = \frac{1}{k}\sum_{j=1}^{k} v_k.$$

(ii) The ***Cesàro means*** for the series $S = \sum_{j=1}^{\infty} v_j$ is the sequence $(\bar{S}^1_k)_{k\in\mathbb{Z}_{>0}}$ of Cesàro means for the sequence of partial sums. Thus

$$\bar{S}^1_k = \frac{1}{k}\sum_{j=1}^{k} S_j = \frac{1}{k}\sum_{j=1}^{k}\sum_{l=1}^{j} v_l.$$

(iii) The sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ is ***Cesàro convergent*** if the sequence $(\bar{v}^1_k)_{k\in\mathbb{Z}_{>0}}$ of Cesàro means converges.

(iv) The series $S = \sum_{j=1}^{\infty} v_j$ is ***Cesàro convergent*** or ***Cesàro summable*** if the sequence $(\bar{S}^1_k)_{k\in\mathbb{Z}_{>0}}$ of Cesàro means converges. •

The us give some examples to illustrate the concept.

6.4.14 Examples (Cesàro convergence)

1. The sequence $(x_j \triangleq (-1)^{j+1})_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{R}$ is oscillatory and so does not converge. However, the sequence is Cesàro convergent since the Cesàro means are given by

$$\bar{x}^1_j = \begin{cases} \frac{1}{j}, & j \text{ odd}, \\ 0, & j \text{ even}, \end{cases}$$

and so the sequence is Cesàro convergent.

2. Let us consider the sum $S = \sum_{j=1}^{\infty}(-1)^{j+1}$ in $\mathbb{R}$. The sequence of partial sums is $(S_k)_{k\in\mathbb{Z}_{>0}}$ with

$$S_k = \begin{cases} 1, & k \text{ odd}, \\ 0, & k \text{ even}. \end{cases}$$

Thus this series is oscillatory. The Cesàro means for the series are $(\bar{S}^1_k)_{k\in\mathbb{Z}_{>0}}$ with

$$\bar{S}^1_k = \begin{cases} \frac{k+1}{2k}, & k \text{ odd}, \\ \frac{1}{2}, & k \text{ even}. \end{cases}$$

Thus the series is Cesàro convergent and the Cesàro means converge to $\frac{1}{2}$. •

[1] Ernesto Cesàro (1859–1906) was an Italian mathematician who made contributions to analysis, number theory, and differential geometry.

The examples illustrate that when one has a divergent sequence or series, it is possible to have Cesàro convergence. This is a useful property that one would ask of a modified version of convergence. The other natural notion is that it should actually generalise the standard notion of convergence. Thus a convergent sequence should still converge with any modified version of convergence. Cesàro convergence possesses this property.

6.4.15 Theorem (Convergence implies Cesàro convergence) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. If a sequence $(\mathrm{v}_j)_{j\in\mathbb{Z}_{>0}}$ (resp. a series $\sum_{j=1}^{\infty} \mathrm{v}_j$) converges to $\mathrm{v}_0 \in \mathsf{V}$ then the sequence (resp. series) converges to v_0 in the sense of Cesàro convergence.*

Proof Since the statement for series follows, by definition, from the statement for sequences, we only show that a convergent sequence is Cesàro convergent with the same limit.

Define $\bar{v}_k^1 = \frac{1}{k}(v_1 + \cdots + v_k)$. Let $\epsilon \in \mathbb{R}_{>0}$ and take $N_1 \in \mathbb{Z}_{>0}$ such that $\|v_j - v_0\| < \frac{\epsilon}{2}$ for $j \geq N_1$. Also take $N_2 \in \mathbb{Z}_{>0}$ sufficiently large that

$$\tfrac{1}{N_2}(\|v_1\| + \cdots + \|v_{N_1}\| + N_1\|v_0\|) < \tfrac{\epsilon}{2}.$$

Then, for $j \geq \{N_1, N_2\}$, we have

$$\begin{aligned}\|\bar{v}_k^1 - v_0\| &= \left\|\tfrac{1}{k}(v_1 + \cdots + v_k) - v_0\right\| = \tfrac{1}{k}\|(v_1 - v_0) + \cdots + (v_k - v_0)\| \\ &\leq \tfrac{1}{k}\|(v_1 - v_0) + \cdots + (v_{N_1} - v_0)\| + \tfrac{1}{k}\|(v_{N_1+1} - v_0) + \cdots + (v_k - v_0)\| \\ &\leq \tfrac{1}{k}(\|v_1\| + \cdots + \|v_{N_1}\| + N_1\|v_0\|) + \tfrac{1}{k}(\|v_{N_1+1} - v_0\| + \cdots + \|v_k - v_0\|) \\ &\leq \frac{\epsilon}{2} + \frac{k - N_1}{k}\frac{\epsilon}{2} < \frac{\epsilon}{2},\end{aligned}$$

giving the result.***missing stuff*** ■

Note that the Cesàro means for a sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ form a sequence $(\bar{v}_j^1)_{j\in\mathbb{Z}_{>0}}$. If this sequence diverges one can ask whether *its* sequence of Cesàro means converges. That is, we can define

$$\bar{v}_k^2 = \frac{1}{k}\sum_{j=1}^{k} \bar{v}_j^1 = \frac{1}{k}\sum_{j=1}^{k}\frac{1}{j}\sum_{l=1}^{j} v_j,$$

and consider the convergence of the sequence $(\bar{v}_k^2)_{k\in\mathbb{Z}_{>0}}$. This can clearly be iterated any finite number of times. This is interesting, although we shall not consider it here. We refer to the notes in Section 6.4.7 for references.

6.4.6 Series in normed vector spaces with arbitrary index sets

In Section 2.4.7 we presented the notion of a series in $\mathbb{R}$ with an arbitrary index set. Such series were useful in discussion saltus functions. Here we discuss series in normed vector spaces with arbitrary index sets. This will be helpful for us in Section 7.3 when we discuss Hilbert bases in general inner product spaces. In any case, much of the treatment mirrors to some extent that for arbitrary series in $\mathbb{R}$.

Let us begin with the definition.

6.4.16 Definition (Convergence of series with arbitrary index sets) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Let A be an index set, consider a family $(v_a)_{a\in A}$ in V, and denote $S = \sum_{a\in A} v_a$. Let $v_0 \in \mathsf{V}$.

(i) The series S ***converges*** to v_0 if, for any $\epsilon \in \mathbb{R}_{>0}$, there exists a finite set $I \subseteq A$ such that

$$\Big\|\sum_{a\in J} v_a - v_0\Big\| < \epsilon$$

for every finite subset $J \subseteq A$ for which $I \subseteq J$.

(ii) The series S is ***Cauchy*** if, for every $\epsilon \in \mathbb{R}_{>0}$, there exists a finite set $I \subseteq A$ such that

$$\Big\|\sum_{a\in J} v_a\Big\| < \epsilon$$

for every finite subset $J \subseteq A$ for which $J \cap I = \emptyset$. •

We already have one point of difference with the results in Section 2.4.7 in that here we have the notion of Cauchy series. This is because we need to allow for the possibility of sums that seem like they should converge, but do not. The next result is analogous to the fact that convergent sequences are always Cauchy, but Cauchy sequences need not converge, but only generally converge when the normed vector space is complete.

6.4.17 Theorem (Relationship between convergent series and Cauchy series) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. For a series $\mathsf{S} = \sum_{\mathrm{a}\in\mathrm{A}} \mathrm{v}_\mathrm{a}$ the following statements hold:*

(i) if S is convergent then it is Cauchy;

(ii) if V is complete and if S is Cauchy then it is convergent.

Proof (i) Let $\epsilon \in \mathbb{R}_{>0}$ and let $I \subseteq A$ be a finite subset such that

$$\Big\|\sum_{a\in J} v_a - v_0\Big\| < \frac{\epsilon}{2}$$

for every finite subset J for which $I \subseteq J$. Let $K \subseteq A$ be finite and such that $K \cap I = \emptyset$. Then

$$\begin{aligned}\Big\|\sum_{a\in K} v_a\Big\| &= \Big\|\sum_{a\in K} v_a + \Big(\sum_{a\in I} v_a - v_0\Big) - \Big(\sum_{a\in I} v_a - v_0\Big)\Big\| \\ &\leq \Big\|\sum_{a\in K\cup I} v_a - v_0\Big\| + \Big\|\sum_{a\in I} v_a - v_0\Big\| \\ &\leq \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,\end{aligned}$$

as desired.

(ii) Let $k \in \mathbb{Z}_{>0}$ and let $I_k \subseteq A$ be a finite subset such that

$$\Big\|\sum_{a\in J} v_a\Big\| < \frac{1}{k}$$

for every finite subset J for which $J \cap I_k = \emptyset$. Then define

$$u_k = \sum_{a \in I_k} v_a.$$

We claim that the sequence $(u_k)_{k\in\mathbb{Z}_{>0}}$ is Cauchy. Indeed, let $N \in \mathbb{Z}_{>0}$ be such that $\frac{1}{N} < \frac{\epsilon}{2}$. Then, for $j, k \geq N$, we have

$$\begin{aligned}\|u_j - u_k\| &= \Big\|\sum_{a\in I_j} v_a - \sum_{a\in I_k} v_a\Big\| = \Big\|\sum_{a\in I_j - I_k} v_a - \sum_{a\in I_k - I_j} v_a\Big\| \\ &\leq \Big\|\sum_{a\in I_j - I_k} v_a\Big\| + \Big\|\sum_{a\in I_k - I_j} v_a\Big\| = \tfrac{1}{j} + \tfrac{1}{k} < \epsilon,\end{aligned}$$

giving $(u_k)_{k\in\mathbb{Z}_{>0}}$ as a Cauchy sequence. Since V is complete there exists a limit u_0 of $(u_k)_{k\in\mathbb{Z}_{>0}}$. Thus, for $\epsilon \in \mathbb{R}_{>0}$, there exists $N_1 \in \mathbb{Z}_{>0}$ such that $\|u_j - u_0\| < \frac{\epsilon}{2}$ for $j \geq N_1$. If $N_2 = \max\{N_1, \frac{2}{\epsilon}\}$ then

$$\begin{aligned}\Big\|\sum_{a\in J} v_a - u_0\Big\| &= \Big\|\sum_{a\in I_{N_2}} v_a - u_0 + \sum_{a\in J\setminus I_{N_2}} v_a\Big\| \\ &\leq \Big\|\sum_{a\in I_{N_2}} v_a - u_0\Big\| + \Big\|\sum_{a\in J\setminus I_{N_2}} v_a\Big\| \leq \tfrac{\epsilon}{2} + \frac{1}{N_2} < \epsilon,\end{aligned}$$

where J is any finite set for which $I_{N_2} \subseteq J$. Thus S converges to u_0. ■

The theorem illustrates the difference between a convergent series and a Cauchy series. The most important fact is that the two notions are equivalent when V is a Banach space.

Just as with arbitrary sums of real numbers, any convergent arbitrary sum in normed vector space can have only countably many nonzero elements.

6.4.18 Proposition (There are only countably many nonzero terms in a convergent series) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. If $\mathrm{S} = \sum_{\mathrm{a}\in\mathrm{A}} \mathrm{v_a}$ is a convergent series then the set $\{\mathrm{a} \in \mathrm{A} \mid \mathrm{v_a} \neq 0_{\mathsf{V}}\}$ is countable.*

Proof By Theorem 6.4.17, since S converges, for any $k \in \mathbb{Z}_{>0}$ there exists a finite set $I_k \subseteq A$ such that

$$\Big\|\sum_{a\in J} v_a\Big\| < \frac{1}{k}$$

for any finite set J such that $J \cap I_k = \emptyset$. Let $I = \cup_{k\in\mathbb{Z}_{>0}} I_k$ so that I is countable by Proposition **??**. If $a \notin I$ then $a \notin I_k$ for all $k \in \mathbb{Z}_{>0}$, i.e., $\{a\} \cap I_k = \emptyset$ for all $k \in \mathbb{Z}_{>0}$. Therefore, $\|v_a\| < \frac{1}{k}$ for all $k \in \mathbb{Z}_{>0}$ and so $\|v_a\| = 0$. Thus $v_a = 0_{\mathsf{V}}$ for all $a \notin I$. ■

Note that Definition 6.4.16 is not the generalisation of Definition 2.4.31, or at least not obviously. Let us prove that the two definitions are, in fact, consistent.

6.4.19 Proposition (Consistency of two notions of arbitrary sums) *Let A be an index set and let $\mathsf{S} = \sum_{\mathsf{a}\in\mathsf{A}} \mathsf{x}_\mathsf{a}$ be a series in $\mathbb{R}$. This series converges according to Definition 2.4.31 if and only if it converges according to Definition 6.4.16, and in case the series converge, they converge to the same limit.*

Proof It suffices to consider the case when the numbers x_a, $a \in A$, are nonnegative (why?). First suppose that S converges according to Definition 2.4.31. Thus

$$\sup\Big\{\sum_{a\in I} x_a \;\Big|\; I \subseteq A \text{ is finite}\Big\} = L < \infty.$$

Let $\epsilon \in \mathbb{R}_{>0}$ and let $I \subseteq A$ be a finite set such that

$$L - \epsilon \le \sum_{a\in I} x_a \le L.$$

Therefore, for any finite set $J \subseteq A$ for which $I \subseteq J$ it holds that

$$L - \epsilon \le \sum_{a\in I} x_a \le \sum_{a\in J} x_a \le L$$

since the elements in the family $(x_a)_{a\in A}$ are nonnegative. This implies that

$$\Big\|\sum_{a\in J} x_a - L\Big\| < \epsilon$$

for any finite set J for which $I \subseteq J$, giving convergence of S to R in the sense of Definition 6.4.16.

The argument above can be essentially reversed to show that if S converges to L in the sense of Definition 6.4.16 then it converges to L in the sense of Definition 2.4.31. ■

For arbitrary series in $\mathbb{R}$ we saw that convergence amounted to absolute convergence in the case when the index set was $\mathbb{Z}_{>0}$. The same is true for arbitrary series in formed vector spaces. For the following result, recall from Proposition 6.4.4 that limits of unconditionally convergent series are independent of rearrangement.

6.4.20 Theorem (A convergent series with index set $\mathbb{Z}_{>0}$ is unconditionally convergent) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. For a sequence $(\mathrm{v}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ the statements are equivalent:*

(i) the series $\sum_{\mathrm{j}\in\mathbb{Z}_{>0}} \mathrm{v}_\mathrm{j}$ is Cauchy in the sense of Definition 6.4.16;

(ii) the series $\sum_{\mathrm{j}=1}^\infty \mathrm{v}_\mathrm{j}$ is unconditionally Cauchy.

Moreover, for $\mathrm{v}_0 \in \mathsf{V}$, the following statements are also equivalent:

(iii) the series $\sum_{\mathrm{j}\in\mathbb{Z}_{>0}} \mathrm{v}_\mathrm{j}$ converges to v_0;

(iv) the series $\sum_{\mathrm{j}=1}^\infty$ converges unconditionally to v_0.

Proof (i) $\Longrightarrow$ (ii) Let $\epsilon \in \mathbb{R}_{>0}$ and let $I \subseteq \mathbb{Z}_{>0}$ be a finite subset such that

$$\Big\|\sum_{j\in J} v_j\Big\| < \epsilon$$

for any finite set $J \subseteq \mathbb{Z}_{>0}$ for which $J \cap I = \emptyset$. Let $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a bijection and choose $N \in \mathbb{Z}_{>0}$ sufficiently large that $I \subseteq \{\phi(1), \dots, \phi(N)\}$. Then, for $k, l \geq N$ with $l > k$ the set $\{\phi(k+1), \dots, \phi(l)\}$ does not intersect I. Thus

$$\Big\| \sum_{j=k+1}^{l} v_{\phi(j)} \Big\| < \epsilon,$$

showing that the lth partial sum minus the kth partial sum is bounded above in norm by ϵ for any $k, l \geq N$. Thus $\sum_{j=1}^{\infty} v_{\phi(j)}$ is Cauchy.

(ii) $\Longrightarrow$ (i) Suppose that (ii) does not hold. Then there exists $\epsilon \in \mathbb{R}_{>0}$ such that, for any finite set $I \subseteq \mathbb{Z}_{>0}$, there exists a finite set $J \subseteq \mathbb{Z}_{>0}$ with $J \cap I = \emptyset$ and such that

$$\Big\| \sum_{j \in J} v_j \Big\| > \epsilon.$$

Now let $I_1 \subseteq \mathbb{Z}_{>0}$ be finite and let $J_1 \subseteq \mathbb{Z}_{>0}$ be finite with $J_1 \cap I_1 = \emptyset$ and with

$$\Big\| \sum_{j \in J_1} v_j \Big\| > \epsilon.$$

Note that $I_2 = I_1 \cup J_1$ is finite. Thus there exists a finite set $J_2 \subseteq \mathbb{Z}_{>0}$ such that $J_2 \cap I_2 = \emptyset$ and such that

$$\Big\| \sum_{j \in J_2} v_j \Big\| > \epsilon.$$

We can continue in this way to define a sequence $(J_k)_{k \in \mathbb{Z}_{>0}}$ of finite pairwise disjoint subsets of $\mathbb{Z}_{>0}$ with the property that

$$\Big\| \sum_{j \in J_k} v_j \Big\| > \epsilon, \qquad k \in \mathbb{Z}_{>0}.$$

Let us denote $\min J_k = m_k$ and $\max J_k = M_k$. Also denote $J_k = \{j_{k,1}, \dots, j_{k,r_k}\}$. Now let $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a bijection such that

$$\phi(\{m_k, \dots, M_k\}) \subseteq \{m_k, \dots, M_k\}$$

and such that

$$\phi(m_k) = j_{k,1}, \dots, \phi(m_k + r_k - 1) = j_{k,r_k}$$

for each $k \in \mathbb{Z}_{>0}$. Then, for any $k \in \mathbb{Z}_{>0}$ we have

$$\Big\| \sum_{j=m_k}^{m_k+r_k-1} v_{\phi(j)} \Big\| = \Big\| \sum_{j \in J_k} v_j \Big\| > \epsilon.$$

Therefore, no matter how large we choose $N \in \mathbb{Z}_{>0}$, there exists $k, l \geq N$ such that the lth partial sum minus the kth partial sum for the series $\sum_{j=1}^{\infty} v_{\phi(j)}$ is bounded below in norm by ϵ. Thus the series is not Cauchy.

(iii) $\Longrightarrow$ (iv) Suppose that $\sum_{j \in \mathbb{Z}_{>0}} v_j$ converges to v_0 in the sense of Definition 6.4.16 to v_0. Let $\epsilon \in \mathbb{R}_{>0}$ and let $I \subseteq \mathbb{Z}_{>0}$ be a finite set such that

$$\Big\| \sum_{j \in J} v_j - v_0 \Big\| < \epsilon$$

for any finite subset $J \subseteq \mathbb{Z}_{>0}$ for which $I \subseteq J$. Let $\phi\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be a bijection. Choose $N \in \mathbb{Z}_{>0}$ sufficiently large that $I \subseteq \{\phi(1), \dots, \phi(N)\}$ and note that, for $k \geq N$ we have

$$\Big\|\sum_{j=1}^{k} v_{\phi(j)} - v_0\Big\| < \epsilon$$

since $S \subseteq \{\phi(1), \dots, \phi(k)\}$. Thus $\sum_{j=1}^{\infty} v_{\phi(j)}$ converges to v_0.

(iv) $\implies$ (iii) Now suppose that $\sum_{j=1}^{\infty} v_j$ converges unconditionally to v_0. Then $\sum_{j=1}^{\infty} v_j$ is unconditionally Cauchy and so Cauchy in the sense of Definition 6.4.16 by the implication (ii) $\implies$ (i). Let $\epsilon \in \mathbb{R}_{>0}$ and let $I' \subseteq \mathbb{Z}_{>0}$ be a finite subset such that

$$\Big\|\sum_{j\in J'} v_j\Big\| < \frac{\epsilon}{2}$$

for every finite subset $J' \subseteq \mathbb{Z}_{>0}$ for which $J' \cap I' = \emptyset$. Let $N \in \mathbb{Z}_{>0}$ be such that

$$\Big\|\sum_{j=1}^{k} v_j - v_0\Big\| < \frac{\epsilon}{2}$$

for every $k \geq N$ and such that $I' \subseteq N$. Define $I = \{1, \dots, N\}$ and let $J \subseteq \mathbb{Z}_{>0}$ be a finite set such that $I \subseteq J$. Write $J = I \cup J'$ with $J' \cap I = \emptyset$. Note that $J' \cap I' = \emptyset$. Therefore,

$$\Big\|\sum_{j\in J} v_j - v_0\Big\| = \Big\|\sum_{j=1}^{N} v_j - v_0 + \sum_{j\in J'} v_j\Big\| \leq \Big\|\sum_{j=1}^{N} v_j - v_0\Big\| + \Big\|\sum_{j\in J'} v_j\Big\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus $\sum_{j\in\mathbb{Z}_{>0}} v_j$ converges to v_0 in the sense of Definition 6.4.16. ■

6.4.7 Notes

We saw in Section 6.4.5 that revised notions of convergence can be applied to divergent series. The classic book of **GHH:49** discusses divergent series in detail.

Theorem 6.4.8 was first proved by **AD/CAR:50**, and the proof we give follows the original proof in form.

Exercises

6.4.1 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Show that

$$\Big\|\sum_{j=1}^{m} v_j\Big\| \leq \sum_{j=1}^{m} \|v_j\|$$

for any finite family $(v_1, \dots, v_m)$ in V.

6.4.2 In Definition 6.4.1 we defined the notions of “convergent series,” “Cauchy series,” “unconditionally convergent series,” and “unconditionally Cauchy series.” We also defined the notion of “absolutely convergent series.” Why did we not define the notion of “absolutely Cauchy series”?

Section 6.5

Continuous maps between normed vector spaces

As with so many areas of mathematics, for normed vector spaces it is interesting to study maps that preserve the structure, in this case the structure defined by the norm. Normed vector spaces have two facets to their structure: (1) the vector space structure and (2) the topology defined by the norm. Thus the interesting maps to consider are linear *and* continuous. We studied linear maps from an algebraic point of view in Sections **??** and **??**, with particular emphasis on the finite-dimensional setting in Section **??**. Maps between topological spaces were the subject of Section **??**. As we shall see, in combining these points of view, one ends up with some quite rich structure.

Do I need to read this section? Continuous linear maps are extremely important in applications. Indeed, the Fourier and Laplace transforms studied in Volume **??** are important examples of continuous linear maps. Therefore, the basic material in this section is important to understand. Some of the more detailed material, for example that in ***missing stuff***, can be skimmed at a first reading, and referred to as needed. •

6.5.1 General continuous maps between normed vector spaces

Most often we will be interested in continuous *linear* maps between normed vector spaces. However, there are also times when it will be helpful to have on hand the notion of continuity for general maps. Thus we present this first.

6.5.1 Definition (Continuous maps between normed vector spaces) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. For open sets $S \subseteq \mathsf{U}$ and $T \subseteq \mathsf{V}$ and for $u_0 \in S$, a map $f\colon S \to T$ is:

(i) ***continuous at*** $\mathbf{u_0}$ if, for each $\epsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that $\|f(u) - f(u_0)\|_{\mathsf{V}} < \epsilon$ whenever $u \in S$ satisfies $\|u - u_0\|_{\mathsf{U}} < \delta$;

(ii) ***continuous*** if it is continuous at each $u_0 \in S$;

(iii) ***uniformly continuous*** if, for each $\epsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that $\|f(u_1) - f(u_2)\| < \epsilon$ for all $u_1, u_2 \in S$ satisfying $\|u_1 - u_2\| < \delta$;

(iv) ***discontinuous at*** $\mathbf{u_0}$ if it is not continuous at u_0;

(v) ***discontinuous*** if it is not continuous. •

We will give interesting examples of continuous *linear* maps in Example 6.5.10. Here let us record some alternative characterisations of continuity.

6.5.2 Theorem (Alternative characterisations of continuity) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *be normed* $\mathbb{F}$*-vector spaces. For a map* $\mathrm{f}\colon \mathrm{S} \to \mathsf{V}$ *defined on an open subset* $\mathrm{S} \subseteq \mathsf{U}$ *and for* $\mathrm{u}_0 \in \mathrm{S}$*, the following statements are equivalent:*

(i) f *is continuous at* u_0*;*

(ii) *for every neighbourhood* B *of* $\mathrm{f}(\mathrm{u}_0)$ *there exists a neighbourhood* A *of* u_0 *in* S *such that* $\mathrm{f}(\mathrm{A}) \subseteq \mathrm{B}$*;*

(iii) $\lim_{\mathrm{u}\to\mathrm{u}_0} \mathrm{f}(\mathrm{u}) = \mathrm{f}(\mathrm{u}_0)$.

Proof In the proof we denote open balls in U and V by $\mathsf{B}_{\mathsf{U}}(r, u)$ and $\mathsf{B}_{\mathsf{V}}(r, v)$, respectively.

(i) $\implies$ (ii) Let $B \subseteq \mathsf{V}$ be a neighbourhood of $f(u_0)$. Let $\epsilon \in \mathbb{R}_{>0}$ be defined such that $\mathsf{B}_{\mathsf{V}}(\epsilon, f(u_0)) \subseteq B$, this being possible since B is open. Since f is continuous at u_0, there exists $\delta \in \mathbb{R}_{>0}$ such that, if $u \in \mathsf{B}_{\mathsf{U}}(\delta, u_0) \cap S$, then we have $f(u) \in \mathsf{B}(\epsilon, f(u_0))$. This shows that, around the point u_0, we can find an open set A in S whose image lies in B.

(ii) $\implies$ (iii) Let $(u_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in S converging to u_0 and let $\epsilon \in \mathbb{R}_{>0}$. By hypothesis there exists a neighbourhood A of u_0 in S such that $f(A) \subseteq \mathsf{B}_{\mathsf{V}}(\epsilon, f(u_0))$. Thus there exists $\delta \in \mathbb{R}_{>0}$ such that $f(\mathsf{B}_{\mathsf{U}}(\delta, u_0) \cap S) \subseteq \mathsf{B}_{\mathsf{V}}(\epsilon, f(u_0))$ since A is open in S. Now choose $N \in \mathbb{Z}_{>0}$ sufficiently large that $|u_j - u_0| < \delta$ for $j \geq N$. It then follows that $|f(u_j) - f(u_0)| < \epsilon$ for $j \geq N$, so giving convergence of $(f(u_j))_{j\in\mathbb{Z}_{>0}}$ to $f(u_0)$, as desired, keeping in mind Notation 6.2.2.

(iii) $\implies$ (i) Let $\epsilon \in \mathbb{R}_{>0}$. Then, by definition of $\lim_{u\to u_0} f(u) = f(u_0)$ from Notation 6.2.2, there exists $\delta \in \mathbb{R}_{>0}$ such that, for $u \in \mathsf{B}_{\mathsf{U}}(\delta, u_0) \cap S$, $|f(u) - f(u_0)| < \epsilon$, which is exactly the definition of continuity of f at u_0. ■

As we have seen, different norms can really be different (i.e., not equivalent), and so, in particular, maps continuous in one norm may not be continuous in another. Moreover, even in finite-dimensions where all norms are equivalent, it is sometimes convenient to use one norm or another, and in this case one would like to ensure that one's conclusions concerning continuity are not dependent on norm. In some sense this is trivial, since equivalent norms define the same topology (Theorem 6.1.14), and it is the topology that determines continuity. However, it is instructive to verify independence of continuity on a choice of equivalent norm. Thus we state the result here, and leave the proof to the reader as Exercise 6.5.2. The result assumes the fact that open sets are the same for equivalent norms; this is exactly what Theorem 6.1.14 shows.

6.5.3 Proposition (Continuity is independent of equivalent norm) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* U *and* V *be* $\mathbb{F}$*-vector spaces, let* $\|\cdot\|_{1,\mathsf{U}}$ *and* $\|\cdot\|_{2,\mathsf{U}}$ *be equivalent norms on* U*, let* $\|\cdot\|_{1,\mathsf{V}}$ *and* $\|\cdot\|_{2,\mathsf{V}}$ *be equivalent norms on* V*, and let* $\mathrm{S} \subseteq \mathsf{U}$ *and* $\mathrm{T} \subseteq \mathsf{V}$ *be open sets. Then, for a map* $\mathrm{f}\colon \mathrm{S} \to \mathrm{T}$*, the following statements are equivalent:*

(i) f *is continuous relative to the norms* $\|\cdot\|_{1,\mathsf{U}}$ *on* U *and* $\|\cdot\|_{1,\mathsf{V}}$ *on* V*;*

(ii) f *is continuous relative to the norms* $\|\cdot\|_{1,\mathsf{U}}$ *on* U *and* $\|\cdot\|_{2,\mathsf{V}}$ *on* V*;*

(iii) f *is continuous relative to the norms* $\|\cdot\|_{2,\mathsf{U}}$ *on* U *and* $\|\cdot\|_{1,\mathsf{V}}$ *on* V*;*

(iv) f *is continuous relative to the norms* $\|\cdot\|_{2,\mathsf{U}}$ *on* U *and* $\|\cdot\|_{2,\mathsf{V}}$ *on* V*.*

With the definition of continuity, let us prove the continuity of some of the standard vector space operations relative to the norm.

6.5.4 Proposition (Continuity properties of operations on normed vector spaces) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a normed* $\mathbb{F}$*-vector space. Then the following maps are continuous:*

(i) $\mathsf{V} \ni \mathrm{v} \mapsto \mathrm{v} + \mathrm{v}_0 \in \mathsf{V}$ *for* $\mathrm{v}_0 \in \mathsf{V}$;

(ii) $\mathsf{V} \oplus \mathsf{V} \ni (\mathrm{v}_1, \mathrm{v}_2) \mapsto \mathrm{v}_1 + \mathrm{v}_2 \in \mathsf{V}$;

(iii) $\mathsf{V} \ni \mathrm{v} \mapsto \mathrm{a}\mathrm{v} \in \mathsf{V}$ *for* $\mathrm{a} \in \mathbb{F}$;

(iv) $\mathbb{F} \oplus \mathsf{V} \ni (\mathrm{a}, \mathrm{v}) \mapsto \mathrm{a}\mathrm{v} \in \mathsf{V}$;

(v) $\mathsf{V} \ni \mathrm{v} \mapsto \|\mathrm{v}\| \in \mathbb{R}$.

Moreover, the maps in parts (i), (ii), (iii), and (v) are uniformly continuous.

Proof (i) For $\epsilon \in \mathbb{R}_{>0}$ let $\delta = \epsilon$. Let $v, v' \in \mathsf{V}$ satisfy $\|v' - v\| < \delta$. We then have

$$\|(v' + v_0) - (v - v_0)\| = \|v' - v\| < \delta = \epsilon,$$

giving uniform continuity of the stated map.

(ii) Let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta = \epsilon$. Let $(u_1, u_2), (v_1, v_2) \in \mathsf{V} \oplus \mathsf{V}$ satisfy $\|(v_1, v_2) - (u_1, u_2)\| < \delta$, where, by abuse of notation, $\|\cdot\|$ denotes the norm on $\mathsf{V} \oplus \mathsf{V}$. Then we have

$$\|v_1 + v_2 - (u_1 - u_2)\| \leq \|v_1 - u_1\| + \|v_2 - u_2\| = \|(v_1, v_2) - (u_1, u_2)\| < \epsilon,$$

giving uniform continuity of the stated map.

(iii) If $a = 0$ then the map is constant, and so certainly uniformly continuous. If $a \neq 0$, let $\epsilon \in \mathbb{R}_{>0}$ and define $\delta = \frac{\epsilon}{|a|}$. Then, if $\|v - v'\| < \delta$ we have

$$\|av - av'\| = |a|\|v - v'\| < \epsilon,$$

giving uniform continuity as desired.

(iv) Let $\epsilon \in \mathbb{R}_{>0}$ and let $(a_0, v_0) \in \mathbb{F} \oplus \mathsf{V}$. Define

$$\delta = \min\Big\{1, \frac{\epsilon}{2(|a_0| + 1)}, \frac{\epsilon}{2(\|v_0\| + 1)}\Big\}$$

and note that if $\|(a, v) - (a_0, v_0)\| < \delta$ (again we abuse notation and denote by $\|\cdot\|$ the norm on $\mathbb{F} \oplus \mathsf{V}$) then we have

$$|a - a_0| + \|v - v_0\| < \delta$$

which in turn implies that

$$\begin{aligned} &|a - a_0| < 1 \quad \Longrightarrow \quad |a| < |a_0| + 1, \\ &|a - a_0| < \frac{\epsilon}{2(|a_0| + 1)}, \\ &\|v - v_0\| < \frac{\epsilon}{2(\|v_0\| + 1)}. \end{aligned}$$

We then compute, for $\|(a, v) - (a_0, v_0)\| < \delta$,

$$\begin{aligned} \|av - a_0v_0\| &= \|av - av_0 + av_0 - a_0v_0\| = \|a(v - v_0) + (a - a_0)v_0\| \\ &\leq |a|\|v - v_0\| + |a - a_0|\|v_0\| \\ &\leq (|a_0| + 1)\frac{\epsilon}{2(|a_0| + 1)} + \frac{\epsilon}{2(\|v_0\| + 1)}(\|v_0\| + 1) = \epsilon. \end{aligned}$$

(v) For $\epsilon \in \mathbb{R}_{>0}$ define $\delta = \epsilon$. Then, if $v, v' \in \mathsf{V}$ satisfy $\|v - v'\| < \delta$, we have

$$|\|v\| - \|v'\|| \leq \|v - v'\| < \delta = \epsilon,$$

giving uniform continuity of the norm. ■

Particularly interesting are continuous bijections with continuous inverses.

6.5.5 Definition (Homeomorphism) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces, and let $S \subseteq \mathsf{U}$ and $T \subseteq \mathsf{V}$ be open sets. A map $f \colon S \to T$ is a ***homeomorphism*** if f is a continuous bijection with a continuous inverse. •

Let us give some examples of homeomorphisms.

6.5.6 Examples (Homeomorphism)

1. The map $f \colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ defined by $f(x) = \tan(x)$ is a homeomorphism with inverse $f^{-1} = \arctan$.
2. Let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space and let $v_0 \in \mathsf{V}$. The map $v \mapsto v + v_0$ is a homeomorphism of V with itself, and has inverse $v \mapsto v - v_0$.
3. Let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space and let $a \in \mathbb{F} \setminus \{0\}$. The map $v \mapsto av$ is a homeomorphism of V with itself, and has inverse $v \mapsto a^{-1}v$. •

6.5.2 Continuous linear maps between normed vector spaces

For vector spaces the maps that preserve the structure are linear maps. For topological spaces the maps that preserve the structure are continuous maps. Thus is makes sense that for normed vector spaces, as they have both the structure of a vector space and a topological space, the most informative maps to consider are those that are linear and continuous. These have a surprisingly rich structure. In this section we give some of their more elementary properties.

Let us first give the notation we will use for continuous linear maps, along with some other useful concepts that can be attached to a linear map.

6.5.7 Definition (Continuous linear maps between normed vector spaces) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}; \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. The set of continuous linear maps from U to V is denoted by $\mathrm{L}(\mathsf{U}; \mathsf{V})$. A linear map $\mathsf{L} \in \mathrm{Hom}_{\mathbb{F}}(\mathsf{U}; \mathsf{V})$ is:

(i) ***bounded*** if there exists $M \in \mathbb{R}_{>0}$ such that $\|\mathsf{L}(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}}$ for every $u \in \mathsf{U}$;

(ii) ***unbounded*** if it is not bounded;

(iii) ***norm-preserving*** if $\|\mathsf{L}(u)\|_{\mathsf{V}} = \|u\|_{\mathsf{U}}$ for all $u \in \mathsf{U}$;

(iv) an ***isomorphism of normed vector spaces*** if it is an isomorphism of vector spaces and is norm-preserving. •

Note that a homeomorphism of normed vector spaces is not necessarily an isomorphism of normed vector spaces, as can be seen in Exercise 6.5.3.

The following result gives a collection of useful conditions that are equivalent to continuity.

6.5.8 Theorem (Characterisations of continuous linear maps) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}; \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. For $\mathsf{L} \in \mathrm{Hom}_{\mathbb{F}}(\mathsf{U}; \mathsf{V})$ the following conditions are equivalent:*

(i) L *is continuous;*

(ii) L *is continuous at* 0_{U}*;*

(iii) L *is uniformly continuous;*

(iv) L *is bounded.*

Moreover, any of the preceding four conditions implies the following:

(v) $\ker(\mathsf{L})$ *is a closed subspace of* U.

Proof (i) $\implies$ (ii) This is clear.

(ii) $\implies$ (iii) Let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta \in \mathbb{R}_{>0}$ such that $\|\mathsf{L}(u)\|_{\mathsf{V}} < \epsilon$ if $\|u\|_{\mathsf{U}} < \delta$; this is possible since L is linear at 0_{U}. Now let $u_0 \in \mathsf{U}$ and suppose that $\|u - u_0\|_{\mathsf{U}} < \delta$. Then

$$\|\mathsf{L}(u) - \mathsf{L}(u_0)\| = \|\mathsf{L}(u - u_0)\| < \epsilon,$$

which gives uniform continuity, as desired.

(iii) $\implies$ (iv) Since L is uniformly continuous, it is continuous at 0_{U}. Let $M \in \mathbb{R}_{>0}$ be such that if $\|u\|_{\mathsf{U}} < \frac{2}{M}$ then $\|\mathsf{L}(u)\|_{\mathsf{V}} < 1$. Let $u \in \mathsf{U}$ and note that

$$\left\|\frac{u}{M\|u\|_{\mathsf{U}}}\right\|_{\mathsf{U}} < \frac{2}{M} \quad\implies\quad \left\|\frac{\mathsf{L}(u)}{M\|u\|_{\mathsf{U}}}\right\|_{\mathsf{V}} < 1 \quad\implies\quad \|\mathsf{L}(u)\|_{\mathsf{V}} < M\|u\|_{\mathsf{U}}.$$

Thus L is bounded.

(iv) $\implies$ (i) Let $M \in \mathbb{R}_{>0}$ be such that $\|\mathsf{L}(u)\|_{\mathsf{V}} < M\|u\|_{\mathsf{U}}$ for all $u \in \mathsf{U}$. For $\epsilon \in \mathbb{R}_{>0}$ let $\delta = \frac{\epsilon}{M}$. If $u_0 \in \mathsf{U}$ and if $\|u - u_0\|_{\mathsf{U}} < \delta$ we have

$$\|\mathsf{L}(u) - \mathsf{L}(u_0)\|_{\mathsf{V}} = \|\mathsf{L}(u - u_0)\|_{\mathsf{V}} \leq M\|u - u_0\|_{\mathsf{U}} < \epsilon.$$

This gives continuity of L.

(iv) $\implies$ (v) Let $(u_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\ker(\mathsf{L})$ converging to $v \in \mathsf{V}$. Then, since L is bounded,

$$\|\mathsf{L}(v) - \mathsf{L}(u_j)\|_{\mathsf{V}} = \|\mathsf{L}(v - u_j)\|_{\mathsf{V}} \leq M\|v - u_j\|_{\mathsf{U}}.$$

Therefore, if $\epsilon \in \mathbb{R}_{>0}$ we can take $N \in \mathbb{Z}_{>0}$ sufficiently large that $\|v - u_j\|_{\mathsf{U}} < \frac{\epsilon}{M}$, and for $j \geq N$ we have

$$\|\mathsf{L}(v)\|_{\mathsf{V}} = \|\mathsf{L}(v) - \mathsf{L}(u_j)\|_{\mathsf{V}} < \epsilon.$$

Thus $\mathsf{L}(v) = 0_{\mathsf{U}}$ and so $v \in \ker(\mathsf{L})$. Thus $\ker(\mathsf{L})$ is closed by Proposition 6.6.8 below. ■

In finite-dimensions, as is so often the case, things simplify.

6.5.9 Theorem (Linear maps from finite-dimensional spaces are continuous) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *be normed* $\mathbb{F}$*-vector spaces. If* U *is finite-dimensional then* $\mathrm{L}(\mathsf{U};\mathsf{V}) = \mathrm{Hom}_{\mathbb{F}}(\mathsf{U};\mathsf{V})$.

Proof Let $\{e_1, \dots, n\}$ be a basis for U and denote

$$M' = \max\{\|\mathsf{L}(e_1)\|_{\mathsf{V}}, \dots, \|\mathsf{L}(e_n)\|_{\mathsf{V}}\}.$$

Define a norm $\|\cdot\|_{1,\mathsf{U}}$ on U by

$$\|u_1e_1 + \cdots + u_ne_n\| = |u_1| + \cdots + |u_n|.$$

By Theorem 6.1.15 there exists $C \in \mathbb{R}_{>0}$ such that $\|u\|_{1,\mathsf{U}} \leq C\|u\|_{\mathsf{U}}$ for all $u \in \mathsf{U}$. Take $M = CM'$. Then, for $u = u_1e_1 + \cdots + u_ne_n \in \mathsf{U}$,

$$\begin{aligned}\|\mathsf{L}(u)\|_{\mathsf{V}} &= \|\mathsf{L}(u_1e_1 + \cdots + u_ne_n)\|_{\mathsf{V}}\\ &\leq |u_1|\|\mathsf{L}(e_1)\|_{\mathsf{V}} + \cdots + |u_n|\|\mathsf{L}(e_n)\|_{\mathsf{V}}\\ &\leq M'\|u\|_{1,\mathsf{U}} \leq M\|u\|_{\mathsf{U}},\end{aligned}$$

showing that L is bounded, and so continuous. ■

Let us give some examples of continuous and discontinuous linear maps, noting that the only interesting examples are infinite-dimensional.

6.5.10 Examples (Continuous linear maps)

1. We take the normed $\mathbb{F}$-vector space $\mathsf{C}^0([a,b];\mathbb{F})$ of continuous $\mathbb{F}$-valued functions on $[a,b]$ equipped with the norm $\|\cdot\|_\infty$ as in Example 6.1.3–10. Define $\mathsf{L}\colon \mathsf{C}^0([a,b];\mathbb{F}) \to \mathsf{C}^0([a,b];\mathbb{F})$ by
$$\mathsf{L}(f)(x) = \int_a^x f(\xi)\,\mathrm{d}\xi.$$
It is easy to show that L is linear, using linearity of the integral. We claim that L is also continuous. To prove this, it suffices to prove that L is continuous at zero. Let $\epsilon \in \mathbb{R}_{>0}$ and let $\delta = \frac{\epsilon}{b-a}$. Then, if $\|f\|_\infty < \delta$,
$$\begin{aligned}\|\mathsf{L}(f)\|_\infty &= \sup\{|\mathsf{L}(f)(x)| \mid x \in [a,b]\}\\ &= \sup\Big\{\Big|\int_a^x f(\xi)\,\mathrm{d}\xi\Big| \;\Big|\; x \in [a,b]\Big\}\\ &\le \sup\Big\{\int_a^x |f(\xi)|\,\mathrm{d}\xi \;\Big|\; x \in [a,b]\Big\}\\ &\le \delta(b-a) = \epsilon,\end{aligned}$$
as desired.
2. Let $\mathsf{C}^1([0,1];\mathbb{R})$ be the $\mathbb{R}$-vector space of continuously differentiable $\mathbb{R}$-valued functions on $[0,1]$. Define $\mathsf{L}\colon \mathsf{C}^1([0,1];\mathbb{R}) \to \mathsf{C}^0([0,1];\mathbb{R})$ by $\mathsf{L}(f) = f'$. By linearity of the derivative, L is linear. We claim that L is not continuous if we use the norm $\|\cdot\|_\infty$ on both $\mathsf{C}^1([0,1];\mathbb{R})$ and $\mathsf{C}^0([0,1];\mathbb{R})$. To show this we shall use the following lemma that is useful in its own right.

1 Lemma *Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$, let $(\mathsf{U};\|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V},\|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces, and let $\mathsf{L} \in \mathrm{Hom}_{\mathbb{F}}(\mathsf{U};\mathsf{V})$. Then L is discontinuous if and only if there exists a sequence $(\mathsf{u}_\mathsf{j})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ in $\mathsf{B}_{\mathsf{U}}(1,0_{\mathsf{U}})$ such that the sequence $(\|\mathsf{L}(\mathsf{u}_\mathsf{j})\|_{\mathsf{V}})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ diverges.*

Proof Suppose that L is continuous. Then there exists $M \in \mathbb{R}_{>0}$ such that $\mathsf{L}(\mathsf{B}_{\mathsf{U}}(1,0_{\mathsf{U}})) \subseteq \mathsf{B}_{\mathsf{V}}(M,0_{\mathsf{V}})$ by boundedness of L. Thus, there can exist no sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{B}_{\mathsf{U}}(1,0_{\mathsf{U}})$ such that the sequence $(\|\mathsf{L}(u_j)\|_{\mathsf{V}})_{j\in\mathbb{Z}_{>0}}$ is unbounded.

No suppose that there is a sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{B}_{\mathsf{U}}(1,0_{\mathsf{U}})$ such that the sequence $(\|\mathsf{L}(u_j)\|_{\mathsf{V}})_{j\in\mathbb{Z}_{>0}}$ diverges. Then, for any $M \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that
$$\|\mathsf{L}(u_j)\|_{\mathsf{V}} \ge M \ge M\|u_j\|_{\mathsf{U}}, \qquad j \ge N.$$
Thus L is unbounded, and so not continuous. ▼

Now consider the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{C}^1([0,1];\mathbb{R})$ given by $f_j(x) = x^j$. This sequence satisfies $\|f_j\|_\infty = 1$. But $\mathsf{L}(f_j)(x) = jx^{j-1}$, and so $\|\mathsf{L}(f_j)\|_\infty = j$, showing that the sequence $(\|\mathsf{L}(f_j)\|_\infty)_{j\in\mathbb{Z}_{>0}}$ diverges. By the lemma it follows that L is discontinuous. •

As a final basic result, let us show that continuous linear maps extend uniquely to the closure. We have not yet defined closure for normed vector spaces, so if you feel like you need to be reminded about what it is, you may refer ahead to Definition 6.6.7.

6.5.11 Proposition (Extension of continuous linear maps to the closure) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *be normed* $\mathbb{F}$*-vector spaces with* V *complete, and let* $\mathsf{W} \subseteq \mathsf{U}$ *be a subspace for which* $\mathrm{cl}(\mathsf{W}) = \mathsf{U}$*. Then, for* $\mathsf{L} \in \mathsf{L}(\mathsf{W}; \mathsf{V})$ *there exists a unique* $\bar{\mathsf{L}} \in \mathsf{L}(\mathsf{U}; \mathsf{V})$ *such that* $\bar{\mathsf{L}}(\mathrm{w}) = \mathsf{L}(\mathrm{w})$ *for all* $\mathrm{w} \in \mathsf{W}$*.*

Proof We let $u \in \mathsf{U}$ and let $(w_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence with the property that $\lim_{j\to\infty} \|u - w_j\|_{\mathsf{U}} = 0$. We first claim that $(\mathsf{L}(w_j))_{j \in \mathbb{Z}_{>0}}$ is a Cauchy sequence. Let $M \in \mathbb{R}_{>0}$ be such that $\|\mathsf{L}(w)\|_{\mathsf{V}} \leq M\|w\|_{\mathsf{U}}$ for all $w \in \mathsf{W}$. Then

$$\|\mathsf{L}(w_j) - \mathsf{L}(w_k)\|_{\mathsf{V}} = \|\mathsf{L}(w_j - w_k)\|_{\mathsf{V}} \leq M\|w_j - w_k\|_{\mathsf{U}}.$$

Since $(w_j)_{j \in \mathbb{Z}_{>0}}$ converges it is a Cauchy sequence, and so it follows that there exists $N \in \mathbb{Z}_{>0}$ for which $\|w_j - w_k\|_{\mathsf{U}} < \frac{\epsilon}{M}$ for $j, k \geq N$. This gives $\|\mathsf{L}(w_j) - \mathsf{L}(w_k)\|_{\mathsf{V}} < \epsilon$ for $j, k \geq N$, so showing that $(\mathsf{L}(w_j))_{j \in \mathbb{Z}_{>0}}$ is indeed a Cauchy sequence. Since $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ is complete, there exists $\bar{\mathsf{L}}(u) \in \mathsf{V}$ which is the limit of the sequence $(\mathsf{L}(w_j))_{j \in \mathbb{Z}_{>0}}$. Next we claim that this limit is independent of the sequence $(w_j)_{j \in \mathbb{Z}_{>0}}$ in W that converges to $u \in \mathsf{U}$. Thus let $(\tilde{w}_j)_{j \in \mathbb{Z}_{>0}}$ be another sequence in W converging to u. We denote by $\tilde{\mathsf{L}}(u)$ the limit in V of the Cauchy sequence $(\mathsf{L}(\tilde{w}_j))_{j \in \mathbb{Z}_{>0}}$. For $j \in \mathbb{Z}_{>0}$ we have

$$\|w_j - \tilde{w}_j\|_{\mathsf{U}} \leq \|w_j - u\|_{\mathsf{U}} + \|\tilde{w}_j - u\|_{\mathsf{U}},$$

implying that $\lim_{j\to\infty} \|w_j - \tilde{w}_j\|_{\mathsf{U}} = 0$. Therefore

$$\|\bar{\mathsf{L}}(u) - \tilde{\mathsf{L}}(u)\|_{\mathsf{V}} \leq \|\bar{\mathsf{L}}(u) - \mathsf{L}(w_j)\|_{\mathsf{V}} + \|\tilde{\mathsf{L}}(u) - \mathsf{L}(\tilde{w}_j)\|_{\mathsf{V}} + \|\mathsf{L}(\tilde{w}_j) - \mathsf{L}(w_j)\|_{\mathsf{V}}.$$

Taking the limit as $j \to \infty$ we see that $\|\bar{\mathsf{L}}(u) - \tilde{\mathsf{L}}(u)\|_{\mathsf{V}}$ can be made smaller than any positive number, and so must be zero.

This then gives us a well-defined element $\bar{\mathsf{L}}(u)$ associated to each $u \in \mathsf{U}$. We next claim that the assignment $u \mapsto \bar{\mathsf{L}}(u)$ is linear. For $u, \tilde{u} \in \mathsf{U}$ let $(w_j)_{j \in \mathbb{Z}_{>0}}$ and $(\tilde{w}_j)_{j \in \mathbb{Z}_{>0}}$ be sequences in W converging to u and $\tilde{u}$, respectively. Then $(w_j + \tilde{w}_j)_{j \in \mathbb{Z}_{>0}}$ converges to $u + \tilde{u}$ by Proposition 6.2.6. Similarly, $(aw_j)_{j \in \mathbb{Z}_{>0}}$ converges to au for $a \in \mathbb{F}$. Therefore

$$\begin{aligned}
\|\bar{\mathsf{L}}(u) + \bar{\mathsf{L}}(\tilde{u}) - \bar{\mathsf{L}}(u + \tilde{u})\|_{\mathsf{V}} &\leq \|\bar{\mathsf{L}}(u) + \bar{\mathsf{L}}(\tilde{u}) - \mathsf{L}(w_j) - \mathsf{L}(\tilde{w}_j)\|_{\mathsf{V}} \\
&\quad + \|\bar{\mathsf{L}}(u + \tilde{u}) - \mathsf{L}(w_j + \tilde{w}_j)\|_{\mathsf{V}} \\
&\leq \|\bar{\mathsf{L}}(u) - \mathsf{L}(w_j)\|_{\mathsf{V}} + \|\bar{\mathsf{L}}(\tilde{u}) - \mathsf{L}(\tilde{w}_j)\|_{\mathsf{V}} \\
&\quad + \|\bar{\mathsf{L}}(u + \tilde{u}) - \mathsf{L}(w_j + \tilde{w}_j)\|_{\mathsf{V}}.
\end{aligned}$$

Taking the limit as $j \to \infty$ shows that the left hand side must be zero, giving $\bar{\mathsf{L}}(u + \tilde{u}) = \bar{\mathsf{L}}(u) + \bar{\mathsf{L}}(u)$. In an entirely similar way we have

$$\|\bar{\mathsf{L}}(au) - a\bar{\mathsf{L}}(u)\|_{\mathsf{V}} \leq \|\bar{\mathsf{L}}(au) - \mathsf{L}(aw_j)\|_{\mathsf{V}} + \|a\bar{\mathsf{L}}(u) - a\mathsf{L}(w_j)\|_{\mathsf{V}},$$

and taking the limit $j \to \infty$ gives $\bar{\mathsf{L}}(au) - a\bar{\mathsf{L}}(u)$.

Let us now demonstrate the uniqueness of the extension $\bar{\mathsf{L}}$. Suppose that $\tilde{\mathsf{L}} \in \mathsf{L}(\mathsf{U}; \mathsf{V})$ is another continuous linear map with the property that it agrees with L on W. For $u \in \mathsf{U}$ let $(w_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in W converging to u. Then

$$\tilde{\mathsf{L}}(u) = \lim_{j\to\infty} \tilde{\mathsf{L}}(w_j) = \lim_{j\to\infty} \mathsf{L}(w_j) = \bar{\mathsf{L}}(w_j)$$

by continuity of $\tilde{\mathsf{L}}$.

missing stuff Finally we show that the operator norm of $\bar{\mathsf{L}}$ is the same as that of L. Since $\bar{\mathsf{L}}$ and L agree on W we have

$$\|\bar{\mathsf{L}}\|_{\mathsf{U},\mathsf{V}} = \sup_{\|u\|_{\mathsf{U}}=1} \|\mathsf{L}(u)\|_{\mathsf{V}} \geq \sup_{\|w\|_{\mathsf{U}}=1} \|\mathsf{L}(w)\|_{\mathsf{V}} = \|\mathsf{L}\|_{\mathsf{W},\mathsf{V}}.$$

Now we prove the opposite inequality. Let $u \in \mathsf{U}$ and let $(w_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in W converging to u. We then have

$$\|\bar{\mathsf{L}}(u)\|_{\mathsf{V}} = \lim_{j\to\infty} \|\mathsf{L}(w_j)\| \leq \lim_{j\to\infty} \|\mathsf{L}\|_{\mathsf{W},\mathsf{V}} \|w_j\|_{\mathsf{U}} = \|\mathsf{L}\|_{\mathsf{W},\mathsf{V}} \|u\|_{\mathsf{U}}.$$

This gives the desired inequality since this must hold for all $u \in \mathsf{U}$, and so concludes the proof. ■

We also have the following related result.

6.5.12 Proposition (Extension of isomorphisms from dense subspaces) *Let* $(\mathsf{V}, \|\cdot\|)$ *be a Banach space with* W *a dense subspace. Suppose that* $\mathsf{L} \in \mathsf{L}_c(\mathsf{V};\mathsf{V})$ *is a continuous linear map with the property that* $\mathsf{L}|\mathsf{W}$ *is a continuous norm preserving bijection from* W *to itself with* $(\mathsf{L}|\mathsf{W})^{-1}$ *being continuous.*[2] *Then* L *is an isomorphism, and* L^{-1} *is the extension, as defined by Proposition 6.5.11, of* $(\mathsf{L}|\mathsf{W})^{-1}$ *to* V*.*

Proof First we note that by Proposition 6.5.11, $\|\mathsf{L}\|_{\mathsf{V},\mathsf{V}} = \|\mathsf{L}|\mathsf{W}\|_{\mathsf{W},\mathsf{W}}$. We claim that this implies that L is norm-preserving. Indeed, let $v \in \mathsf{V}$ and let $(w_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in W converging to v. Then

$$\|\mathsf{L}(u)\| = \lim_{j\to\infty} \|\mathsf{L}(w_j)\| = \lim_{j\to\infty} \|w_j\| = \|u\|,$$

as desired. We next claim that this implies injectivity of L. Indeed, if $\mathsf{L}(v) = 0$ for $v \in \mathsf{V}$ we must then have $\|v\| = \|\mathsf{L}(v)\| = 0$, giving $v = 0$. Thus L is injective. We also claim that image(L) is a closed subspace. Let $(\mathsf{L}(v_j))_{j\in\mathbb{Z}_{>0}}$ be a sequence in image(L) converging to $u \in \mathsf{V}$. Then since $\|\mathsf{L}(v_j) - \mathsf{L}(v_k)\| = \|v_j - v_k\|$ it follows that $(v_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. Let $v \in \mathsf{V}$ denote the limit of this sequence. We need to show that $\mathsf{L}(v) = u$. Indeed,

$$\|\mathsf{L}(v) - u\| \leq \|\mathsf{L}(v) - \mathsf{L}(v_j)\| + \|u - \mathsf{L}(v_j)\|,$$

and taking the limit as $j \to \infty$ gives $\|\mathsf{L}(v) - u\| = 0$, so showing that image($\mathsf{L}$) is closed. Since $\mathsf{W} \subseteq$ image(L) and since cl(W) = V we must have cl(image(L)) = image(L) = V, thus showing surjectivity of L.

Finally we must show that L^{-1} is the unique continuous extension of $(\mathsf{L}|\mathsf{W})^{-1}$ to V. Let M denote the unique continuous extension of $(\mathsf{L}|\mathsf{W})^{-1}$ to V. Just as L is a continuous bijection, so too is M. Let $v \in \mathsf{V}$ and let $(w_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in W converging to $\mathsf{L}(v)$. Then

$$\mathsf{M} \circ \mathsf{L}(v) = \lim_{j\to\infty} (\mathsf{L}|\mathsf{W})^{-1}(w_j).$$

There then exists a sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$ so that $\mathsf{L}(u_j) = w_j$, $j \in \mathbb{Z}_{>0}$. We then have

$$\mathsf{M} \circ \mathsf{L}(v) = \lim_{j\to\infty} (\mathsf{L}|\mathsf{W})^{-1} \circ \mathsf{L}(u_j) = \lim_{j\to\infty} u_j.$$

[2] The assumption that $(\mathsf{L}|\mathsf{W})^{-1}$ be continuous is actually superfluous by the ***Banach Isomorphism Theorem***.

We claim that $\lim_{j\to\infty} u_j = v$. Since L is continuous and injective, this is equivalent to showing that $\lim_{j\to\infty} \mathsf{L}(u_j) = \mathsf{L}(v)$. However, this follows directly from the definition of the sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$. Next let $v \in \mathsf{V}$ and let $(w_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in W converging to $\mathsf{M}(v)$. Then

$$\mathsf{L}\circ\mathsf{M}(v) = \lim_{j\to\infty} \mathsf{L}(w_j).$$

Let $(u_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in W with the property that $(\mathsf{L}|\mathsf{W})^{-1}(u_j) = w_j$, $j \in \mathbb{Z}_{>0}$. Then we have

$$\mathsf{L}\circ\mathsf{M}(v) = \lim_{j\to\infty} \mathsf{L}\circ(\mathsf{L}|\mathsf{W})^{-1}(u_j) = \lim_{j\to\infty} u_j.$$

We must show that $\lim_{j\to\infty} u_j = v$. Since M is continuous and injective this is equivalent to showing that $\lim_{j\to\infty} \mathsf{M}(u_j) = \mathsf{M}(v)$. This follows, however, from the definition of the sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$. Thus we have shown that $\mathsf{M}\circ\mathsf{L}(v) = \mathsf{L}\circ\mathsf{M}(v) = v$ for all $v \in \mathsf{V}$. Thus $\mathsf{M} = \mathsf{L}^{-1}$. ■

6.5.3 Induced topologies on continuous linear maps

Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. In Corollary **??** we showed that $\mathrm{Hom}_{\mathbb{F}}(\mathsf{U};\mathsf{V})$ is an $\mathbb{F}$-vector space. This is a purely algebraic observation. Now we wish to study the structure of the continuous linear maps. As we shall see, this is itself a normed vector space.

First we should establish that the set of continuous linear maps form a vector space.

6.5.13 Proposition (L(U; V) is a subspace of Hom$_{\mathbb{F}}$(U; V)) *If* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ *and if* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *are normed* $\mathbb{F}$*-vector spaces, then* $\mathrm{L}(\mathsf{U};\mathsf{V})$ *is a subspace of* $\mathrm{Hom}_{\mathbb{F}}(\mathsf{U};\mathsf{V})$.

Proof Let $\mathsf{L}_1, \mathsf{L}_2 \in \mathrm{L}(\mathsf{U};\mathsf{V})$. For $\epsilon \in \mathbb{R}_{>0}$ let $\delta \in \mathbb{R}_{>0}$ be such that $\|\mathsf{L}_1(u)\|_{\mathsf{V}} < \frac{\epsilon}{2}$ and $\|\mathsf{L}_2(u)\|_{\mathsf{V}} < \frac{\epsilon}{2}$ for $\|u\|_{\mathsf{U}} < \delta$. Then compute

$$\|(\mathsf{L}_1 + \mathsf{L}_2)(u)\|_{\mathsf{V}} \le \|\mathsf{L}_1(u)\|_{\mathsf{V}} + \|\mathsf{L}_2(u)\|_{\mathsf{V}} < \epsilon,$$

showing that $\mathsf{L}_1 + \mathsf{L}_2$ is continuous at 0_{U}, and so continuous. Also let $a \in \mathbb{F}$ and $\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V})$. If $a = 0$ it is clear that $a\mathsf{L}$ is continuous. So suppose that $a \neq 0$, let $\epsilon \in \mathbb{R}_{>0}$, and let $\delta \in \mathbb{R}_{>0}$ be such that if $\|u\|_{\mathsf{U}} < \delta$ then $\|\mathsf{L}(u)\|_{\mathsf{V}} < \frac{\epsilon}{|a|}$. For $\|u\|_{\mathsf{U}} < \delta$ we then have

$$\|(a\mathsf{L})(u)\|_{\mathsf{V}} = |a|\|\mathsf{L}(u)\|_{\mathsf{V}} < \epsilon,$$

giving continuity of $a\mathsf{L}$. ■

This shows that $\mathrm{L}(\mathsf{U};\mathsf{V})$ is indeed an $\mathbb{F}$-vector space. It is moreover true that it is a *normed* vector space.

6.5.14 Theorem (L(U; V) is a normed vector space) *Let* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ *and let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *be normed* $\mathbb{F}$*-vector spaces. For* $\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V})$ *define*

$$\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = \inf\{\mathrm{M} \in \mathbb{R}_{>0} \mid \|\mathsf{L}(\mathrm{u})\|_{\mathsf{V}} \le \mathrm{M}\|\mathrm{u}\|_{\mathsf{U}},\ \mathrm{u} \in \mathsf{U}\}.$$

Then $\|\cdot\|_{\mathsf{U},\mathsf{V}}$ *is a norm on* $\mathrm{L}(\mathsf{U};\mathsf{V})$*. Moreover,*

(i) $\|\mathsf{L}(\mathrm{u})\|_{\mathsf{V}} \le \|\mathsf{L}\|_{\mathsf{U},\mathsf{V}}\|\mathrm{u}\|_{\mathsf{U}}$ *for all* $\mathrm{u} \in \mathsf{U}$,

(ii) $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = \sup\left\{\frac{\|\mathsf{L}(\mathrm{u})\|_{\mathsf{V}}}{\|\mathrm{u}\|_{\mathsf{U}}} \;\middle|\; \mathrm{u} \in \mathsf{U} \setminus \{0_{\mathsf{V}}\}\right\}$,

(iii) $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = \sup\{\|\mathsf{L}(\mathrm{u})\|_{\mathsf{V}} \mid \|\mathrm{u}\|_{\mathsf{U}} = 1\}$, *and*

(iv) $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = \sup\{\|\mathsf{L}(\mathrm{u})\|_{\mathsf{V}} \mid \|\mathrm{u}\|_{\mathsf{U}} \leq 1\}$, *and*

(v) *if* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *is complete then so is* $(\mathrm{L}(\mathsf{U};\mathsf{V}), \|\cdot\|_{\mathsf{U},\mathsf{V}})$.

Proof Let us first verify (i), disregarding whether or not $\|\cdot\|_{\mathsf{U},\mathsf{V}}$ is a norm. Suppose that (i) does not hold. Then there exists $u \in \mathsf{U}$ such that $\|\mathsf{L}(u)\|_{\mathsf{V}} > \|\mathsf{L}\|_{\mathsf{U},\mathsf{V}}\|u\|_{\mathsf{U}}$. Thus there exists $\epsilon \in \mathbb{R}_{>0}$ such that

$$\|\mathsf{L}(u)\|_{\mathsf{V}} > (\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} - \epsilon)\|u\|_{\mathsf{U}},$$

and this contradicts the definition of $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}}$.

We next note that $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} \in \mathbb{R}_{>0}$ for every $\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V})$. Moreover, $\|0_{\mathrm{L}(\mathsf{U};\mathsf{V})}\|_{\mathsf{U},\mathsf{V}} = 0$. Now suppose that $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = 0$. Then

$$\|\mathsf{L}(u)\|_{\mathsf{V}} \leq \|\mathsf{L}\|_{\mathsf{U},\mathsf{V}}\|u\|_{\mathsf{U}} = 0, \qquad u \in \mathsf{U}.$$

Thus $\mathsf{L} = 0_{\mathrm{L}(\mathsf{U};\mathsf{V})}$. Clearly we have $\|0\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = |0|\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}}$. If $a \in \mathbb{F} \setminus \{0\}$ then we compute

$$\begin{aligned}
\|a\mathsf{L}\|_{\mathsf{U},\mathsf{V}} &= \inf\{M \in \mathbb{R}_{>0} \mid \|a\mathsf{L}(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}}, u \in \mathsf{U}\} \\
&= \inf\{M \in \mathbb{R}_{>0} \mid |a|\|\mathsf{L}(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}}, u \in \mathsf{U}\} \\
&= \inf\left\{M \in \mathbb{R}_{>0} \;\middle|\; \|\mathsf{L}(u)\|_{\mathsf{V}} \leq \frac{M}{|a|}\|u\|_{\mathsf{U}}, u \in \mathsf{U}\right\} \\
&= \inf\{|a|M' \in \mathbb{R}_{>0} \mid \|\mathsf{L}(u)\|_{\mathsf{V}} \leq M'\|u\|_{\mathsf{U}}, u \in \mathsf{U}\} = |a|\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}},
\end{aligned}$$

using Proposition 2.2.28. Finally, if $\mathsf{L}_1, \mathsf{L}_2 \in \mathrm{L}(\mathsf{U};\mathsf{V})$ then

$$\begin{aligned}
\|\mathsf{L}_1 + \mathsf{L}_2\|_{\mathsf{U},\mathsf{V}} &= \inf\{M \in \mathbb{R}_{>0} \mid \|(\mathsf{L}_1 + \mathsf{L}_2)(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}},\ u \in \mathsf{U}\} \\
&\leq \inf\{M \in \mathbb{R}_{>0} \mid \|\mathsf{L}_1(u)\|_{\mathsf{V}} + \|\mathsf{L}_2(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}},\ u \in \mathsf{U}\} \\
&= \inf\{M_1 + M_2 \in \mathbb{R}_{>0} \mid \|\mathsf{L}_1(u)\|_{\mathsf{V}} \leq M_1\|u\|_{\mathsf{U}}, \\
&\quad \|\mathsf{L}_2(u)\|_{\mathsf{V}} \leq M_2\|u\|_{\mathsf{U}},\ u \in \mathsf{U}\} \\
&= \inf\{M \in \mathbb{R}_{>0} \mid \|\mathsf{L}_1(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}},\ u \in \mathsf{U}\} \\
&\quad + \inf\{M \in \mathbb{R}_{>0} \mid \|\mathsf{L}_2(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}},\ u \in \mathsf{U}\} \\
&= \|\mathsf{L}_1\|_{\mathsf{U},\mathsf{V}} + \|\mathsf{L}_2\|_{\mathsf{U},\mathsf{V}},
\end{aligned}$$

where we have used Proposition 2.2.28. This verifies that $\|\cdot\|_{\mathsf{U},\mathsf{V}}$ has the properties demanded of a norm.

(ii) First note that the equality is trivial when $\mathsf{L} = 0_{\mathrm{L}(\mathsf{U};\mathsf{V})}$, so we suppose this is not the case. In this case, $\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} > 0$ and so

$$\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} = \inf\{M \in \mathbb{R}_{>0} \mid \|\mathsf{L}(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}},\ u \in \mathsf{U} \setminus \{0_{\mathsf{V}}\}\}$$

and so

$$\begin{aligned}
\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} &= \inf\{M \in \mathbb{R}_{>0} \mid \|\mathsf{L}(u)\|_{\mathsf{V}} \leq M\|u\|_{\mathsf{U}},\ u \in \mathsf{U} \setminus \{0_{\mathsf{V}}\}\} \\
&= \inf\left\{M \in \mathbb{R}_{>0} \;\middle|\; \frac{\|\mathsf{L}(u)\|_{\mathsf{V}}}{\|u\|_{\mathsf{U}}} \leq M,\ u \in \mathsf{U} \setminus \{0_{\mathsf{U}}\}\right\} \\
&= \sup\left\{\frac{\|\mathsf{L}(u)\|_{\mathsf{V}}}{\|u\|_{\mathsf{U}}} \;\middle|\; u \in \mathsf{U} \setminus \{0_{\mathsf{U}}\}\right\}.
\end{aligned}$$

(iii) Carrying on from part (ii) we have

$$\begin{aligned}\|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} &= \sup\Big\{\frac{\|\mathsf{L}(u)\|_{\mathsf{V}}}{\|u\|_{\mathsf{U}}}\ \Big|\ u \in \mathsf{U} \setminus \{0_{\mathsf{U}}\}\Big\} \\ &= \sup\Big\{\Big\|\mathsf{L}\Big(\frac{u}{\|u\|_{\mathsf{U}}}\Big)\Big\|\ \Big|\ u \in \mathsf{U} \setminus \{0_{\mathsf{U}}\}\Big\} \\ &= \sup\{\|\mathsf{L}(u)\|_{\mathsf{V}} \mid \|u\|_{\mathsf{U}} = 1\}.\end{aligned}$$

(iv) It is evident that

$$\sup\{\|\mathsf{L}(u)\|_{\mathsf{V}} \mid \|u\|_{\mathsf{U}} \leq 1\} \geq \sup\{\|\mathsf{L}(u)\|_{\mathsf{V}} \mid \|u\|_{\mathsf{U}} = 1\},$$

the supremum on the left being taken over a larger set. On the other hand,

$$\begin{aligned}\sup\{\|\mathsf{L}(u)\|_{\mathsf{V}} \mid \|u\|_{\mathsf{U}} \leq 1\} &= \sup\{\|\mathsf{L}(\lambda u)\|_{\mathsf{V}} \mid \lambda \in [0,1],\ \|u\|_{\mathsf{U}} = 1\} \\ &= \sup\{\lambda\|\mathsf{L}(u)\|_{\mathsf{V}} \mid \lambda \in [0,1],\ \|u\|_{\mathsf{U}} = 1\} \\ &\leq \sup\{\|\mathsf{L}(u)\|_{\mathsf{V}} \mid \|u\|_{\mathsf{U}} = 1\},\end{aligned}$$

giving the result.

(v) Let $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathrm{L}(\mathsf{U};\mathsf{V})$. We claim that $(\mathsf{L}_j(u))_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in V. This is clear if $u = 0_{\mathsf{U}}$, so let us suppose otherwise. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that $\|\mathsf{L}_j - \mathsf{L}_k\|_{\mathsf{U},\mathsf{V}} < \frac{\epsilon}{\|u\|_{\mathsf{U}}}$ for $j,k \geq N$. Then

$$\|\mathsf{L}_j(u) - \mathsf{L}_k(u)\|_{\mathsf{V}} \leq \|\mathsf{L}_j - \mathsf{L}_k\|_{\mathsf{U},\mathsf{V}}\|u\|_{\mathsf{V}} < \epsilon$$

for $j,k \geq N$. Thus the sequence $(\mathsf{L}_j(u))_{j\in\mathbb{Z}_{>0}}$ converges to an element in V which we denote by $\mathsf{L}(u)$. One may easily show that the assignment $u \mapsto \mathsf{L}(u)$ is well-defined and linear, cf. the proof of Proposition 6.5.11. Thus this defines $\mathsf{L} \in \mathrm{Hom}_{\mathbb{F}}(\mathsf{U};\mathsf{V})$.

We now show that L is continuous. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that $\|\mathsf{L}_j - \mathsf{L}_k\|_{\mathsf{U},\mathsf{V}} < \epsilon$ for $j,k \geq N$. Then, if $\|u\|_{\mathsf{U}} \leq 1$,

$$\|(\mathsf{L}_j - \mathsf{L}_k)(u)\|_{\mathsf{V}} \leq \|\mathsf{L}_j - \mathsf{L}_k\|_{\mathsf{U},\mathsf{V}}\|u\|_{\mathsf{U}} < \epsilon.$$

Using continuity of the norm and Theorem 6.5.2 we have, for fixed $j \geq N$,

$$\lim_{k\to\infty}\|(\mathsf{L}_j - \mathsf{L}_k)(u)\|_{\mathsf{V}} = \Big\|(\mathsf{L}_j - \lim_{k\to\infty}\mathsf{L}_k)(u)\Big\|_{\mathsf{V}} = \|(\mathsf{L}_j - \mathsf{L})(u)\|_{\mathsf{V}} < \epsilon.$$

Therefore, for any $u \in \mathsf{U}$ we have

$$\|(\mathsf{L}_j - \mathsf{L})(u)\|_{\mathsf{V}} < \epsilon\|u\|_{\mathsf{U}},$$

implying that $\mathsf{L}_j - \mathsf{L}$ is bounded and so $\mathsf{L}_j - \mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V})$. Since $\mathsf{L}_j \in \mathrm{L}(\mathsf{U};\mathsf{V})$ and since $\mathrm{L}(\mathsf{U};\mathsf{V})$ is a subspace it follows that $\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V})$.

Moreover, our computations also show that, for any $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that $\|\mathsf{L}_j - \mathsf{L}\|_{\mathsf{U},\mathsf{V}} < \epsilon$ for $j \geq N$. Thus $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges to L. ■

Let us attach some terminology to our norm on $\mathrm{L}(\mathsf{U};\mathsf{V})$.

6.5.15 Definition (Induced norm, operator norm, convergence in norm) Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $(\mathsf{U},\|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V},\|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces.

(i) The norm $\|\cdot\|_{\mathsf{U},\mathsf{V}}$ is the ***induced norm*** or the ***operator norm*** on $\mathrm{L}(\mathsf{U};\mathsf{V})$.

(ii) A sequence $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ ***converges in norm*** if it converges in the normed $\mathbb{F}$-vector space $(\mathrm{L}(\mathsf{U};\mathsf{V}),\|\cdot\|_{\mathsf{U},\mathsf{V}})$. •

The induced norm also satisfies nice properties with respect to composition.

6.5.16 Proposition (Induced norm and composition) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$, $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$, *and* $(\mathsf{W}, \|\cdot\|_{\mathsf{W}})$ *be normed* $\mathbb{F}$*-vector spaces. If* $\mathsf{L} \in \mathrm{L}(\mathsf{U}; \mathsf{V})$ *and* $\mathsf{K} \in \mathrm{L}(\mathsf{V}; \mathsf{W})$ *then*

$$\|\mathsf{K} \circ \mathsf{L}\|_{\mathsf{U},\mathsf{W}} \le \|\mathsf{K}\|_{\mathsf{V},\mathsf{W}} \|\mathsf{L}\|_{\mathsf{U},\mathsf{V}}.$$

In particular, $\mathsf{K} \circ \mathsf{L} \in \mathrm{L}(\mathsf{U}; \mathsf{W})$.

Proof For $u \in \mathsf{U}$ we compute

$$\|\mathsf{K} \circ \mathsf{L}(u)\|_{\mathsf{W}} \le \|\mathsf{K}\|_{\mathsf{V},\mathsf{W}} \mathsf{L}(u) \le \|\mathsf{K}\|_{\mathsf{V},\mathsf{W}} \|\mathsf{L}\|_{\mathsf{U},\mathsf{V}} \|u\|_{\mathsf{U}},$$

as desired. ■

As suggested by the terminology "converges in norm," we wish to allow other versions of convergence of sequences of continuous linear maps. The principal such notion is the following.

6.5.17 Definition (Strong convergence) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. A sequence $(\mathsf{L}_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathrm{L}(\mathsf{U}; \mathsf{V})$ ***converges strongly*** to $\mathsf{L} \in \mathrm{L}(\mathsf{U}; \mathsf{V})$ if, for each $u \in \mathsf{U}$, the sequence $(\mathsf{L}_j(u))_{j \in \mathbb{Z}_{>0}}$ converges. •

Let us explore strong convergence by providing its relationship with convergence in norm.

6.5.18 Proposition (Convergence in norm implies strong convergence) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *be normed* $\mathbb{F}$*-vector spaces. A sequence* $(\mathsf{L}_{\mathsf{j}})_{\mathsf{j} \in \mathbb{Z}_{>0}}$ *in* $\mathrm{L}(\mathsf{U}; \mathsf{V})$ *converges strongly if it converges in norm.*

Proof This is Exercise 6.5.4. ■

It is not generally true that strong convergence implies convergence in norm. The following example relies on the reader knowing about Banach spaces of sequences as discussed in Section 6.7.2.

6.5.19 Example (Strong convergence may not imply norm convergence) We consider the $\mathbb{F}$-Banach space $\ell^2(\mathbb{F})$ of sequences $(a_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathbb{F}$ for which $\sum_{j=1}^{\infty} |a_j|^2 < \infty$. This is a Banach space with norm

$$\|(a_j)_{j \in \mathbb{Z}_{>0}}\|_2 = \Big(\sum_{j=1}^{\infty} |a_j|^2\Big)^{1/2}.$$

For $k \in \mathbb{Z}_{>0}$ define $\mathsf{L}_k \in \mathrm{L}(\ell^2(\mathbb{R}); \mathbb{F})$ by $\mathsf{L}_k((a_j)_{j \in \mathbb{Z}_{>0}}) = a_k$ (it is clear that L_k is linear and bounded). Now note that

$$(\mathsf{L}_k - \mathsf{L}_l)((a_j)_{j \in \mathbb{Z}_{>0}}) = a_k - a_l$$

so that

$$|(\mathsf{L}_k - \mathsf{L}_l)((a_j)_{j \in \mathbb{Z}_{>0}})| \le |a_k| + |a_l| \le \sqrt{2}(|a_k|^2 + |a_l|^2)^{1/2} \le \sqrt{2} \|(a_j)_{j \in \mathbb{Z}_{>0}}\|_2,$$

where we have used Proposition **??**. Thus $\|\mathsf{L}_k - \mathsf{L}_l\|_{\ell^2(\mathbb{F}),\mathbb{F}} \leq \sqrt{2}$. However, taking the particular sequence

$$a_j = \begin{cases} 1, & j = k, \\ -1, & j = l, \\ 0, & \text{otherwise}, \end{cases}$$

we have

$$|(\mathsf{L}_k - \mathsf{L}_l)((a_j)_{j\in\mathbb{Z}_{>0}})| = \sqrt{2}\|(a_j)_{j\in\mathbb{Z}_{>0}}\|_2,$$

showing that $\|\mathsf{L}_k - \mathsf{L}_l\|_{\ell^2(\mathbb{F}),\mathbb{F}} \leq \sqrt{2}$. In particular, the sequence $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ is not Cauchy, and so does not converge in norm. We claim that it does, however, converge strongly. Indeed, if $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^2(\mathbb{F})$ then we have $\lim_{j\to\infty}|a_j|^2 = 0$ by Proposition 2.4.7. Therefore,

$$\lim_{k\to\infty} \mathsf{L}_k((a_j)_{j\in\mathbb{Z}_{>0}}) = \lim_{k\to\infty} a_k = 0,$$

showing that the sequence $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges strongly to the zero linear map. •

The preceding example notwithstanding, the reader may not be surprised to learn that strong and norm convergence agree in finite-dimensions.

6.5.20 Proposition (Equivalence of strong and norm convergence in finite-dimensions) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be finite-dimensional normed $\mathbb{F}$-vector spaces. A sequence $(\mathsf{L}_{\mathsf{j}})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ in $\mathsf{L}(\mathsf{U};\mathsf{V})$ converges strongly if and only if it converges in norm.*

Proof Let $\{e_1, \dots, e_n\}$ be a basis for U. We claim that

$$|||\mathsf{L}||| = \max\{\|\mathsf{L}(e_1)\|_{\mathsf{V}}, \dots, \|\mathsf{L}(e_n)\|_{\mathsf{V}}\}$$

is a norm on $\mathsf{L}(\mathsf{U};\mathsf{V})$. The only possibly nontrivial fact to verify is the triangle inequality. For this we have

$$\begin{aligned} |||\mathsf{L}_1 + \mathsf{L}_2||| &= \max\{\|(\mathsf{L}_1 + \mathsf{L}_2)(e_1)\|_{\mathsf{V}}, \dots, \|(\mathsf{L}_1 + \mathsf{L}_2)(e_n)\|_{\mathsf{V}}\} \\ &\leq \max\{\|\mathsf{L}_1(e_1)\|_{\mathsf{V}} + \|\mathsf{L}_2(e_1)\|_{\mathsf{V}}, \dots, \|\mathsf{L}_1(e_n)\|_{\mathsf{V}} + \|\mathsf{L}_2(e_n)\|_{\mathsf{V}}\} \\ &= \max\{\|\mathsf{L}_1(e_1)\|_{\mathsf{V}}, \dots, \|\mathsf{L}_1(e_n)\|_{\mathsf{V}}\} + \max\{\|\mathsf{L}_2(e_1)\|_{\mathsf{V}}, \dots, \|\mathsf{L}_2(e_n)\|_{\mathsf{V}}\} \\ &= |||\mathsf{L}_1||| + |||\mathsf{L}_2|||, \end{aligned}$$

as desired.

Now we claim that a sequence $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{L}(\mathsf{U};\mathsf{V})$ converges strongly to L if and only if it converges to L in the norm $|||\cdot|||$. Indeed, strong convergence implies immediately that $\lim_{j\to\infty} \mathsf{L}_j(e_k) = \mathsf{L}(e_k)$ for each $k \in \{1, \dots, n\}$. This in turn implies convergence in the norm $|||\cdot|||$. Conversely, if a sequence $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges in the norm $|||\cdot|||$ then, for each $k \in \{1, \dots, n\}$, $(\mathsf{L}_j(e_k))_{j\in\mathbb{Z}_{>0}}$ converges in V to $\mathsf{L}(e_k)$. Thus, if $u = u_1e_1 + \cdots + u_ne_n \in \mathsf{U}$ we have,

$$\lim_{j\to\infty} \mathsf{L}_j(u_1e_1 + \cdots + u_ne_n) = \sum_{k=1}^{n} u_k \lim_{j\to\infty} \mathsf{L}(e_k) = \mathsf{L}(u_1e_1 + \cdots + u_ne_n).$$

Thus $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges to L strongly.

The result follows from this since the norms $|||\cdot|||$ and $\|\cdot\|_{\mathsf{U},\mathsf{V}}$ are equivalent by virtue of $\mathsf{L}(\mathsf{U};\mathsf{V})$ being finite-dimensional (see Exercise **??**). ■

We close this section by indicating that strong convergence is, in fact, convergence in a suitable topology. The material here relies on an understanding of topics covered in ***missing stuff***. It is not necessary to understand this to understand strong convergence.

6.5.21 Definition (Strong operator topology) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. The ***strong operator topology*** is the topology for which sets of the form

$$\cap_{k=1}^{m}\{\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V}) \mid \|\mathsf{L}(u_k) - \mathsf{L}_0(u_k)\| < \epsilon_k\}, \quad u_1, \dots, u_m \in \mathsf{U},\ \epsilon_1, \dots, \epsilon_m \in \mathbb{R}_{>0},$$

are a neighbourhood basis about L_0. •

That this does indeed define a topology on $\mathrm{L}(\mathsf{U};\mathsf{V})$ follows from ***missing stuff***.

The following result connects the strong operator topology with the notion of strong convergence.

6.5.22 Theorem (Strong convergence is convergence in the strong operator topology) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces. Then a sequence $(\mathsf{L}_{\mathsf{j}})_{\mathsf{j}\in\mathbb{Z}_{>0}}$ in $\mathrm{L}(\mathsf{U};\mathsf{V})$ converges strongly to L_0 if and only if it converges to L_0 in the strong operator topology.*

Proof First suppose that $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges strongly to L_0. Let $S \subseteq \mathrm{L}(\mathsf{U};\mathsf{V})$ be a neighbourhood of L_0 in the strong operator topology and let $\epsilon_1, \dots, \epsilon_k \in \mathbb{R}_{>0}$ and $u_1, \dots, u_k \in \mathsf{U}$ be such that

$$\cap_{k=1}^{m}\{\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V}) \mid \|\mathsf{L}(u_k) - \mathsf{L}_0(u_k)\|_{\mathsf{V}} < \epsilon_k\} \subseteq S.$$

For $k \in \{1, \dots, m\}$ let $N_k \in \mathbb{Z}_{>0}$ be sufficiently large that $\|\mathsf{L}_j(u_k) - \mathsf{L}_0(u_k)\|_{\mathsf{V}} < \epsilon_k$ for $j \geq N_k$ and let $N = \max\{N_1, \dots, N_m\}$. Then, for $j \geq N$ and for $k \in \{1, \dots, m\}$,

$$\|\mathsf{L}_j(u_k) - \mathsf{L}_0(u_k)\|_{\mathsf{V}} < \epsilon_k$$

so that

$$\mathsf{L}_j \in \cap_{k=1}^{m}\{\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V}) \mid \|\mathsf{L}(u_k) - \mathsf{L}_0(u_k)\|_{\mathsf{V}} < \epsilon_k\}.$$

Thus $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges in the strong operator topology.

Now suppose that $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ converges to L_0 in the strong operator topology. For $\epsilon \in \mathbb{R}_{>0}$ and $u \in \mathsf{U}$ note that

$$S(\mathsf{L}_0, u, \epsilon) \triangleq \{\mathsf{L} \in \mathrm{L}(\mathsf{U};\mathsf{V}) \mid \|\mathsf{L}(u) - \mathsf{L}_0(u)\|_{\mathsf{V}} < \epsilon\}$$

is a neighbourhood of L_0 in the strong operator topology. Thus, for $\epsilon \in \mathbb{R}_{>0}$ and $u \in \mathsf{U}$, there exists $N \in \mathbb{Z}_{>0}$ such that $\mathsf{L}_j \in S(\mathsf{L}_0, u, \epsilon)$ for $j \geq N$. That is, for each $\epsilon \in \mathbb{R}_{>0}$ and for each $u \in \mathsf{U}$, there exists $N \in \mathbb{Z}_{>0}$ such that $\|\mathsf{L}(u) - \mathsf{L}_0(u)\|_{\mathsf{V}} < \epsilon$ showing that $(\mathsf{L}_j(u))_{j\in\mathbb{Z}_{>0}}$ converges to $\mathsf{L}_0(u)$. This is exactly strong convergence of $(\mathsf{L}_j)_{j\in\mathbb{Z}_{>0}}$ to L_0. ■

6.5.23 Remark (The strong operator topology is locally convex) As a glimpse ahead to Chapter **??** we make the observation that the strong operator topology is the locally convex topology defined by the family of seminorms $(p_u)_{u\in\mathsf{U}}$ where $p_u(\mathsf{L}) = \|\mathsf{L}(u)\|_{\mathsf{V}}$. •

missing stuff

6.5.4 The Open Mapping Theorem and Closed Graph Theorem

In the preceding two***missing stuff*** sections we studied some of the more basic characterisations of continuous linear maps between normed vector spaces. In the next ***missing stuff*** sections we give some deeper results which provide some very useful structure for Banach spaces.

6.5.24 Theorem (Banach–Schauder Open Mapping Theorem) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ *and* $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ *be* $\mathbb{F}$*-Banach spaces. If* $\mathsf{L} \in \mathsf{L}(\mathsf{U};\mathsf{V})$ *is surjective then it is open, i.e.,* $\mathsf{L}(\mathsf{S})$ *is open for every open subset* $\mathsf{S} \subseteq \mathsf{U}$*.*

Proof ■

It is worth reflecting on whether it is necessary that U and V be Banach spaces in order for the result to hold. It turns out that these assumptions are necessary.

6.5.25 Examples (Open Mapping Theorem fails for normed vector spaces)

1. Consider the following data:
   (a) $\mathsf{U} = \mathsf{C}^0([0,1],\mathbb{R})$;
   (b) $\|\cdot\|_{\mathsf{U}} = \|\cdot\|_\infty$;
   (c) V is the subspace of functions f in $\mathsf{C}^0([0,1];\mathbb{R})$ that are continuously differentiable and that satisfy $f(0) = 0$;
   (d) $\|\cdot\|_{\mathsf{V}} = \|\cdot\|_\infty$;
   (e) $\mathsf{L} \in \mathrm{Hom}_{\mathbb{R}}(\mathsf{U};\mathsf{V})$ is defined by
   $$\mathsf{L}(f)(x) = \int_0^x f(\xi)\,\mathrm{d}\xi.$$

   Note that V is not complete; we invite the reader to adapt Example 6.6.25–2 to provide a Cauchy sequence in V that does not converge.

   We claim that L is a continuous bijection but its inverse is not continuous.
2. Let $(\mathsf{V}, \|\cdot\|)$ be a Banach space of infinite-dimension and let $\{e_i\}_{i\in I}$ be a basis for V, and suppose without loss of generality that $\|e_i\| = 1$ for each $i \in I$. As in the proof of Proposition 6.1.4 define a norm $\|\cdot\|_1$ on V by
   $$\Big\|\sum_{i\in I} c_i e_i\Big\|_1 = \sum_{i\in I} |c_i|,$$
   this definition making sense since the sum is finite. As in ***missing stuff***, $(\mathsf{V}, \|\cdot\|_1)$ is incomplete. We claim that the identity map on V, thought of as a linear map from the normed vector space $(\mathsf{V}, \|\cdot\|_1)$ to the Banach space $(\mathsf{V}, \|\cdot\|)$, is a continuous bijection but has an inverse that is not continuous.

Exercises

6.5.1 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ and $(\mathsf{V}, \|\cdot\|_{\mathsf{V}})$ be normed $\mathbb{F}$-vector spaces, and let $\mathsf{L} \in \mathrm{L}(\mathsf{U}; \mathsf{V})$. Show that L is norm-preserving if and only if it is an isometry of the metric spaces associated with the norms (cf. Proposition 6.1.7).

6.5.2 Prove Proposition 6.5.3.

6.5.3 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. On $\mathbb{F}^n$ consider the two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ as in Example 6.1.3. Show that $\mathrm{id}_{\mathbb{F}^n}$ is a homeomorphism of the normed vector spaces $(\mathbb{F}^n, \|\cdot\|_1)$ and $(\mathbb{F}^n, \|\cdot\|_2)$ but is not an isomorphism of normed vector spaces.

6.5.4 Prove Proposition 6.5.18.

Section 6.6

Topology of normed vector spaces

Since a (semi)normed vector space is a metric space, and so a topological space, one has all of the usual notions associated with topological spaces: interior, closure, boundary, compactness, etc. These notions inherit all of the attributes from general topological spaces as discussed in detail in Chapter **??**. We would like, however, for the reader to be able to at least read the results in this section without having first read Chapter **??**. Therefore, we adopt the following approach for presentation. All definitions and theorems are stated so that they can be read independently of having read Chapter **??**. When it is easily done, proofs are given in a way that does not rely on understanding general notions from topology. However, we also do not shy away from using some general ideas from Chapter **??** in a proof when doing so avoids duplication. The bottom line is this: A reader should be able to understand the flow of ideas without having read Chapter **??**, but understanding all proofs may require understanding some parts of Chapter **??**.

It is also the case that, like quite a few of the results in this chapter, the statements and proofs bear a strong resemblance to those for real numbers; the reader should thus compare what we say here with what has been said already in Section 2.5. The similarities and the differences together will help reader understand normed vector spaces.

Do I need to read this section? Readers already familiar with topology can forgo the basic definitions and theorems.***missing stuff*** The notion of a Schauder basis in Section 6.6.5 will come up in ***missing stuff***. •

6.6.1 Properties of balls in normed vector spaces

In this section we give some fairly easy and pretty "obvious" results concerning the character of open and closed balls in normed vector spaces. These results will be used constantly in our description of the topology of normed vector spaces.

We know that, by definition, the open balls in a normed vector space form a basis for the norm topology; every open set is by definition a union of open balls. This description can be refined a little to show that it is really open balls about 0_{V} that are important.

6.6.1 Proposition (Balls about the origin are sufficient to describe the norm topology) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a normed* $\mathbb{F}$*-vector space. For any open set* $\mathsf{U} \subseteq \mathsf{V}$ *there exists an index set* I*, positive numbers* $(r_i)_{i\in I}$*, vectors* $(v_i)_{i\in I}$ *such that*

$$\mathsf{U} = \cup_{i\in I}\{v_i + \mathsf{B}(r_i, 0_{\mathsf{V}})\},$$

where

$$v_i + \mathsf{B}(r_i, 0_{\mathsf{V}}) = \{v + v_i \mid v \in \mathsf{B}(r_i, 0_{\mathsf{V}})\}.$$

Proof This follows since $\mathsf{B}(r, v)$ is the translation by v of $\mathsf{B}(r, 0_{\mathsf{V}})$, cf. the proof of Proposition 6.1.12. ■

Let us next give some fairly elementary properties of open and closed balls.

6.6.2 Proposition (Properties of open and closed balls) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space, and let $\mathrm{r} \in \mathbb{R}_{>0}$ and $\mathrm{v}_0 \in \mathsf{V}$. Then the following statements hold:*

(i) $\mathsf{B}(\mathrm{r}, \mathrm{v}_0)$ *is open;*

(ii) $\overline{\mathsf{B}}(\mathrm{r}, \mathrm{v}_0)$ *is closed and bounded;*

(iii) $\overline{\mathsf{B}}(\mathrm{r}, \mathrm{v}_0)$ *is compact if and only if* $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ *is compact.*

Proof (i) This is Exercise 6.1.1.

(ii) If $M = \|v\| + r$ and if $v \in \overline{\mathsf{B}}(r, v)$ then

$$\|v\| = \|v - v_0 + v_0\| \leq \|v - v_0\| + \|v_0\| \leq M,$$

showing that $\overline{\mathsf{B}}(r, v_0) \subseteq \overline{\mathsf{B}}(M, 0_{\mathsf{V}})$ and so $\overline{\mathsf{B}}(r, v_0)$ is bounded. Define $f\colon \mathsf{V} \to \mathbb{R}$ by $f(v) = \|v\|$ and note that $\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) = f^{-1}([0, 1])$. Since f is continuous by Proposition 6.5.4 and since $[0, 1]$ is closed, it follows that $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is closed by Proposition **??**. Now define $f_r, f_{v_0}\colon \mathsf{V} \to \mathsf{V}$ by $f_r(v) = rv$ and $f_{v_0}(v) = v + v_0$. By Proposition 6.5.4 these maps are homeomorphisms. Therefore, $f_{v_0} \circ f_r$ is continuous. Since $\overline{\mathsf{B}}(r, v_0) = f_{v_0} \circ f_r(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}))$ and since the homeomorphic image of a closed set is closed (Corollary **??**), it follows that $\overline{\mathsf{B}}(r, v_0)$ is closed.

(iii) As in the preceding part of the proof, $\overline{\mathsf{B}}(r, v_0) = f_{v_0} \circ f_r(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}))$, and since the continuous image of compact sets is compact (Proposition **??**), the result follows. ■

6.6.2 Interior, closure, boundary, etc.

The definitions and results here are similar to those for $\mathbb{R}$ given in Section 2.5.3, so we will go through them quickly. Examples, discussion, and motivation can be found in Section 2.5.3.

6.6.3 Definition (Neighbourhood) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. For $v \in \mathsf{V}$, a ***neighbourhood*** of v is an open set U for which $v \in U$. •

6.6.4 Definition (Accumulation point, cluster point, limit point) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. For a subset $A \subseteq \mathsf{V}$, a point $v \in \mathsf{V}$ is:

(i) an ***accumulation point*** for A if, for every neighbourhood U of v, the set $A \cap (U \setminus \{v\})$ is nonempty;

(ii) a ***cluster point*** for A if, for every neighbourhood U of v, the set $A \cap U$ is infinite;

(iii) a ***limit point*** of A if there exists a sequence $(v_j)_{j \in \mathbb{Z}_{>0}}$ in A converging to v.

The set of accumulation points of A is called the ***derived set*** of A, and is denoted by $\operatorname{der}(A)$. •

In Remark 2.5.12 we made some comments about conventions concerning the words “accumulation point,” “cluster point,” and “limit point.” Those remarks apply equally here.

6.6.5 Proposition ("Accumulation point" equals "cluster point") *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a normed* $\mathbb{F}$*-vector space. For a set* $\mathrm{A} \subseteq \mathsf{V}$, $\mathrm{v} \in \mathsf{V}$ *is an accumulation point for* A *if and only if it is a cluster point for* A.

Proof It is clear that a cluster point for A is an accumulation point for A. Suppose that v is not a cluster point. Then there exists a neighbourhood U of v for which the set $A \cap U$ is finite. If $A \cap U = \{v\}$, then clearly v is not an accumulation point. If $A \cap U \neq \{v\}$, then $A \cap (U \setminus \{v\}) \supseteq \{v_1, \dots, v_k\}$ where the points $v_1, \dots, v_k$ are distinct from v. Now let

$$\epsilon = \tfrac{1}{2} \min\{\|v_1 - v\|, \dots, \|v_k - v\|\}.$$

Clearly $A \cap (\mathsf{B}(\epsilon, v) \setminus \{v\})$ is then empty, and so v is not an accumulation point for A. ■

6.6.6 Proposition (Properties of the derived set) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a normed* $\mathbb{F}$*-vector space. For* $\mathrm{A}, \mathrm{B} \subseteq \mathsf{V}$ *and for a family of subsets* $(\mathrm{A}_i)_{i \in I}$ *of* V*, the following statements hold:*

(i) $\mathrm{der}(\emptyset) = \emptyset$*;*

(ii) $\mathrm{der}(\mathsf{V}) = \mathsf{V}$*;*

(iii) $\mathrm{der}(\mathrm{der}(\mathrm{A})) = \mathrm{der}(\mathrm{A})$*;*

(iv) if $\mathrm{A} \subseteq \mathrm{B}$ *then* $\mathrm{der}(\mathrm{A}) \subseteq \mathrm{der}(\mathrm{B})$*;*

(v) $\mathrm{der}(\mathrm{A} \cup \mathrm{B}) = \mathrm{der}(\mathrm{A}) \cup \mathrm{der}(\mathrm{B})$*;*

(vi) $\mathrm{der}(\mathrm{A} \cap \mathrm{B}) \subseteq \mathrm{der}(\mathrm{A}) \cap \mathrm{der}(\mathrm{B})$.

Proof Parts (i) and (ii) follow directly from the definition of the derived set.

(iii) ***missing stuff***

(iv) Let $v \in \mathrm{der}(A)$ and let U be a neighbourhood of v. Then the set $A \cap (U \setminus \{v\})$ is nonempty, implying that the set $B \cap (U \setminus \{v\})$ is also nonempty. Thus $v \in \mathrm{der}(B)$.

(v) Let $v \in \mathrm{der}(A \cup B)$ and let U be a neighbourhood of v. Then the set $U \cap ((A \cup B) \setminus \{v\})$ is nonempty. But

$$\begin{aligned} U \cap ((A \cup B) \setminus \{v\}) &= U \cap ((A \setminus \{v\}) \cup (B \setminus \{v\})) \\ &= (U \cap (A \setminus \{v\})) \cup (U \cap (B \setminus \{v\})). \end{aligned} \tag{6.6}$$

Thus it cannot be that both $U \cap (A \setminus \{v\})$ and $U \cap (B \setminus \{v\})$ are empty. Thus x is an element of either $\mathrm{der}(A)$ or $\mathrm{der}(B)$.

Now let $v \in \mathrm{der}(A) \cup \mathrm{der}(A)$. Then, using (6.6), $U \cap ((A \cup B) \setminus \{v\})$ is nonempty, and so $v \in \mathrm{der}(A \cup B)$.

(vi) Let $x \in \mathrm{der}(A \cap B)$ and let U be a neighbourhood of v. Then $U \cap ((A \cap B) \setminus \{v\}) \neq \emptyset$. We have

$$U \cap ((A \cap B) \setminus \{v\}) = U \cap ((A \setminus \{v\}) \cap (B \setminus \{v\}))$$

Thus the sets $U \cap (A \setminus \{v\})$ and $U \cap (B \setminus \{v\})$ are both nonempty, showing that $v \in \mathrm{der}(A) \cap \mathrm{der}(B)$. ■

6.6.7 Definition (Interior, closure, and boundary) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. Let $A \subseteq \mathsf{V}$.

(i) The ***interior*** of A is the set

$$\operatorname{int}(A) = \cup\{U \mid U \subseteq A,\ U \text{ open}\}.$$

(ii) The ***closure*** of A is the set

$$\operatorname{cl}(A) = \cap\{C \mid A \subseteq C,\ C \text{ closed}\}.$$

(iii) The ***boundary*** of A is the set $\operatorname{bd}(A) = \operatorname{cl}(A) \cap \operatorname{cl}(\mathsf{V} \setminus A)$. •

6.6.8 Proposition (Characterisation of interior, closure, and boundary) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a normed* $\mathbb{F}$*-vector space. For* $A \subseteq \mathsf{V}$*, the following statements hold:*

(i) $v \in \operatorname{int}(A)$ *if and only if there exists a neighbourhood* U *of* v *such that* $U \subseteq A$*;*

(ii) $v \in \operatorname{cl}(A)$ *if and only if, for each neighbourhood* U *of* v*, the set* $U \cap A$ *is nonempty;*

(iii) $v \in \operatorname{bd}(A)$ *if and only if, for each neighbourhood* U *of* v*, the sets* $U \cap A$ *and* $U \cap (\mathsf{V} \setminus A)$ *are nonempty.*

Proof (i) Suppose that $v \in \operatorname{int}(A)$. Since $\operatorname{int}(A)$ is open, there exists a neighbourhood U of v contained in $\operatorname{int}(A)$. Since $\operatorname{int}(A) \subseteq A$, $U \subseteq A$.

Next suppose that $v \notin \operatorname{int}(A)$. Then, by definition of interior, for any open set U for which $U \subseteq A$, $v \notin U$.

(ii) Suppose that there exists a neighbourhood U of v such that $U \cap A = \emptyset$. Then $\mathsf{V} \setminus U$ is a closed set containing A. Thus $\operatorname{cl}(A) \subseteq \mathsf{V} \setminus U$. Since $v \notin \mathsf{V} \setminus U$, it follows that $v \notin \operatorname{cl}(A)$.

Suppose that $v \notin \operatorname{cl}(A)$. Then v is an element of the open set $\mathsf{V} \setminus \operatorname{cl}(A)$. Thus there exists a neighbourhood U of v such that $U \subseteq \mathsf{V} \setminus \operatorname{cl}(A)$. In particular, $U \cap A = \emptyset$.

(iii) This follows directly from part (ii) and the definition of boundary. ■

6.6.9 Proposition (Properties of interior) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a normed* $\mathbb{F}$*-vector space. For* $A, B \subseteq \mathsf{V}$ *and for a family of subsets* $(A_i)_{i \in I}$ *of* V*, the following statements hold:*

(i) $\operatorname{int}(\emptyset) = \emptyset$*;*

(ii) $\operatorname{int}(\mathsf{V}) = \mathsf{V}$*;*

(iii) $\operatorname{int}(\operatorname{int}(A)) = \operatorname{int}(A)$*;*

(iv) if $A \subseteq B$ *then* $\operatorname{int}(A) \subseteq \operatorname{int}(B)$*;*

(v) $\operatorname{int}(A \cup B) \supseteq \operatorname{int}(A) \cup \operatorname{int}(B)$*;*

(vi) $\operatorname{int}(A \cap B) = \operatorname{int}(A) \cap \operatorname{int}(B)$*;*

(vii) $\operatorname{int}(\cup_{i \in I} A_i) \supseteq \cup_{i \in I} \operatorname{int}(A_i)$*;*

(viii) $\operatorname{int}(\cap_{i \in I} A_i) \subseteq \cap_{i \in I} \operatorname{int}(A_i)$*.*

Moreover, a set $A \subseteq \mathsf{V}$ *is open if and only if* $\operatorname{int}(A) = A$*.*

Proof Parts (i) and (ii) are clear by definition of interior. Part (v) follows from part (vii), so we will only prove the latter.

(iii) This follows since the interior of an open set is the set itself.

(iv) Let $v \in \mathrm{int}(A)$. Then there exists a neighbourhood U of v such that $U \subseteq A$. Thus $U \subseteq B$, and the result follows from Proposition 6.6.8.

(vi) Let $v \in \mathrm{int}(A) \cap \mathrm{int}(B)$. Since $\mathrm{int}(A) \cap \mathrm{int}(B)$ is open by Exercise 2.5.1, there exists a neighbourhood U of v such that $U \subseteq \mathrm{int}(A) \cap \mathrm{int}(B)$. Thus $U \subseteq A \cap B$. This shows that $v \in \mathrm{int}(A \cap B)$. This part of the result follows from part (viii).

(vii) Let $v \in \cup_{i\in I} \mathrm{int}(A_i)$. By Exercise 2.5.1 the set $\cup_{i\in I} \mathrm{int}(A_i)$ is open. Thus there exists a neighbourhood U of v such that $U \subseteq \cup_{i\in I} \mathrm{int}(A_i)$. Thus $U \subseteq \cup_{i\in I} A_i$, from which we conclude that $v \in \mathrm{int}(\cup_{i\in I} A_i)$.

(viii) Let $v \in \mathrm{int}(\cap_{i\in I} A_i)$. Then there exists a neighbourhood U of v such that $U \subseteq \cap_{i\in I} A_i$. It therefore follows that $U \subseteq A_i$ for each $i \in I$, and so that $v \in \mathrm{int}(A_i)$ for each $i \in I$.

The final assertion follows directly from Proposition 6.6.8. ■

6.6.10 Proposition (Properties of closure) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. For $\mathrm{A}, \mathrm{B} \subseteq \mathsf{V}$ and for a family of subsets $(\mathrm{A}_i)_{i\in I}$ of V, the following statements hold:*

(i) $\mathrm{cl}(\emptyset) = \emptyset$;

(ii) $\mathrm{cl}(\mathsf{V}) = \mathsf{V}$;

(iii) $\mathrm{cl}(\mathrm{cl}(\mathrm{A})) = \mathrm{cl}(\mathrm{A})$;

(iv) *if* $\mathrm{A} \subseteq \mathrm{B}$ *then* $\mathrm{cl}(\mathrm{A}) \subseteq \mathrm{cl}(\mathrm{B})$;

(v) $\mathrm{cl}(\mathrm{A} \cup \mathrm{B}) = \mathrm{cl}(\mathrm{A}) \cup \mathrm{cl}(\mathrm{B})$;

(vi) $\mathrm{cl}(\mathrm{A} \cap \mathrm{B}) \subseteq \mathrm{cl}(\mathrm{A}) \cap \mathrm{cl}(\mathrm{B})$;

(vii) $\mathrm{cl}(\cup_{i\in I} \mathrm{A}_i) \supseteq \cup_{i\in I} \mathrm{cl}(\mathrm{A}_i)$;

(viii) $\mathrm{cl}(\cap_{i\in I} \mathrm{A}_i) \subseteq \cap_{i\in I} \mathrm{cl}(\mathrm{A}_i)$.

Moreover, a set $\mathrm{A} \subseteq \mathsf{V}$ is closed if and only if $\mathrm{cl}(\mathrm{A}) = \mathrm{A}$.

Proof Parts (i) and (ii) follow immediately from the definition of closure. Part (vi) follows from part (viii), so we will only prove the latter.

(iii) This follows since the closure of a closed set is the set itself.

(iv) Suppose that $v \in \mathrm{cl}(A)$. Then, for any neighbourhood U of v, the set $U \cap A$ is nonempty, by Proposition 6.6.8. Since $A \subseteq B$, it follows that $U \cap B$ is also nonempty, and so $v \in \mathrm{cl}(B)$.

(v) Let $v \in \mathrm{cl}(A \cup B)$. Then, for any neighbourhood U of v, the set $U \cap (A \cup B)$ is nonempty by Proposition 6.6.8. By Proposition 1.1.4, $U \cap (A \cup B) = (U \cap A) \cup (U \cap B)$. Thus the sets $U \cap A$ and $U \cap B$ are not both nonempty, and so $v \in \mathrm{cl}(A) \cup \mathrm{cl}(B)$. That $\mathrm{cl}(A) \cup \mathrm{cl}(B) \subseteq \mathrm{cl}(A \cup B)$ follows from part (vii).

(vi) Let $v \in \mathrm{cl}(A \cap B)$. Then, for any neighbourhood U of v, the set $U \cap (A \cap B)$ is nonempty. Thus the sets $U \cap A$ and $U \cap B$ are nonempty, and so $v \in \mathrm{cl}(A) \cap \mathrm{cl}(B)$.

(vii) Let $v \in \cup_{i\in I} \mathrm{cl}(A_i)$ and let U be a neighbourhood of v. Then, for each $i \in I$, $U \cap A_i \neq \emptyset$. Therefore, $\cup_{i\in I}(U \cap A_i) \neq \emptyset$. By Proposition 1.1.7, $\cup_{i\in I}(U \cap A_i) = U \cap (\cup_{i\in I} A_i)$, showing that $U \cap (\cup_{i\in I} A_i) \neq \emptyset$. Thus $v \in \mathrm{cl}(\cup_{i\in I} A_i)$.

(viii) Let $v \in \mathrm{cl}(\cap_{i\in I} A_i)$ and let U be a neighbourhood of v. Then the set $U \cap (\cap_{i\in I} A_i)$ is nonempty. This means that, for each $i \in I$, the set $U \cap A_i$ is nonempty. Thus $v \in \mathrm{cl}(A_i)$ for each $i \in I$, giving the result. ■

6.6.11 Proposition (Joint properties of interior, closure, boundary, and derived set) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. For $\mathrm{A} \subseteq \mathsf{V}$, the following statements hold:*

(i) $\mathsf{V} \setminus \mathrm{int}(\mathrm{A}) = \mathrm{cl}(\mathsf{V} \setminus \mathrm{A})$*;*

(ii) $\mathsf{V} \setminus \mathrm{cl}(\mathrm{A}) = \mathrm{int}(\mathsf{V} \setminus \mathrm{A})$.

(iii) $\mathrm{cl}(\mathrm{A}) = \mathrm{A} \cup \mathrm{bd}(\mathrm{A})$*;*

(iv) $\mathrm{int}(\mathrm{A}) = \mathrm{A} - \mathrm{bd}(\mathrm{A})$*;*

(v) $\mathrm{cl}(\mathrm{A}) = \mathrm{int}(\mathrm{A}) \cup \mathrm{bd}(\mathrm{A})$*;*

(vi) $\mathrm{cl}(\mathrm{A}) = \mathrm{A} \cup \mathrm{der}(\mathrm{A})$*;*

(vii) $\mathsf{V} = \mathrm{int}(\mathrm{A}) \cup \mathrm{bd}(\mathrm{A}) \cup \mathrm{int}(\mathsf{V} \setminus \mathrm{A})$.

Proof (i) Let $v \in \mathsf{V} \setminus \mathrm{int}(A)$. Since $v \notin \mathrm{int}(A)$, for every neighbourhood U of v it holds that $U \not\subseteq A$. Thus, for any neighbourhood U of v, we have $U \cap (\mathsf{V} \setminus A) \neq \emptyset$, showing that $v \in \mathrm{cl}(\mathsf{V} \setminus A)$.

Now let $v \in \mathrm{cl}(\mathsf{V} \setminus A)$. Then for any neighbourhood U of v we have $U \cap (\mathsf{V} \setminus A) \neq \emptyset$. Thus $v \notin \mathrm{int}(A)$, so $v \in \mathsf{V} \setminus A$.

(ii) The proof here strongly resembles that for part (i), and we encourage the reader to provide the explicit arguments.

(iii) This follows from part (v).

(iv) Clearly $\int(A) \subseteq A$. Suppose that $v \in A \cap \mathrm{bd}(A)$. Then, for any neighbourhood U of v, the set $U \cap (\mathsf{V} \setminus A)$ is nonempty. Therefore, no neighbourhood of v is a subset of A, and so $v \notin \mathrm{int}(A)$. Conversely, if $v \in \mathrm{int}(A)$ then there is a neighbourhood U of v such that $U \subseteq A$. The precludes the set $U \cap (\mathsf{V} \setminus A)$ from being nonempty, and so we must have $v \notin \mathrm{bd}(A)$.

(v) Let $v \in \mathrm{cl}(A)$. For a neighbourhood U of v it then holds that $U \cap A \neq \emptyset$. If there exists a neighbourhood V of v such that $V \subseteq A$, then $v \in \mathrm{int}(A)$. If there exists *no* neighbourhood V of v such that $V \subseteq A$, then for every neighbourhood V of v we have $V \cap (\mathsf{V} \setminus A) \neq \emptyset$, and so $v \in \mathrm{bd}(A)$.

Now let $v \in \mathrm{int}(A) \cup \mathrm{bd}(A)$. If $v \in \mathrm{int}(A)$ then $v \in A$ and so $v \in\subseteq \mathrm{cl}(A)$. If $v \in \mathrm{bd}(A)$ then it follows immediately from Proposition 6.6.8 that $v \in \mathrm{cl}(A)$.

(vi) Let $v \in \mathrm{cl}(A)$. If $v \notin A$ then, for every neighbourhood U of v, $U \cap A = U \cap (A \setminus \{v\}) \neq \emptyset$, and so $v \in \mathrm{der}(A)$.

If $v \in A \cup \mathrm{der}(A)$ then either $v \in A \subseteq \mathrm{cl}(A)$, or $v \notin A$. In this latter case, $v \in \mathrm{der}(A)$ and so the set $U \cap (A \setminus \{v\})$ is nonempty for each neighbourhood U of v, and we again conclude that $v \in \mathrm{cl}(A)$.

(vii) Clearly $\mathrm{int}(A) \cap \mathrm{int}(\mathsf{V} \setminus A) = \emptyset$ since $A \cap (\mathsf{V} \setminus A) = \emptyset$. Now let $v \in \mathsf{V} \setminus (\mathrm{int}(A) \cup \mathrm{int}(\mathsf{V} \setminus A))$. Then, for any neighbourhood U of v, we have $U \not\subseteq A$ and $U \not\subseteq (\mathsf{V} \setminus A)$. Thus the sets $U \cap (\mathsf{V} \setminus A)$ and $U \cap A$ must both be nonempty, from which we conclude that $v \in \mathrm{bd}(A)$. ■

Let us close this section with a discussion of some notions not present in Section 2.5, but which are important for normed vector spaces. General topological versions of these ideas have been discussed in ***missing stuff***.

6.6.12 Definition (Dense, nowhere dense, separable) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. Let $A, B \subseteq \mathsf{V}$ with $A \subseteq B$.

(i) The set A is ***dense*** in B if $\mathrm{cl}(A) = B$.

(ii) The set A is ***nowhere dense*** in B if $B \setminus \mathrm{cl}(A)$ is dense in B.

(iii) The set A is ***separable*** if there exists a countable dense subset of A. •

We refer to ***missing stuff*** for simple examples that illustrate these definitions. Generally speaking, it is not uncommon to see the requirement that a Banach space be separable, although there are important examples of nonseparable Banach spaces, as we shall see in Section 6.7.

6.6.3 Compactness

As we shall shortly see, the discussion of compactness for normed vector spaces has a different flavour than that for compact subsets of $\mathbb{R}$. This is because compactness in infinite-dimensional normed vector spaces is quite a strict notion, for example more strict than closed and bounded. However, the initial definitions proceed just as for $\mathbb{R}$.

We begin with simple definitions concerning open covers.

6.6.13 Definition (Open cover) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space, and let $A \subseteq \mathsf{V}$.

(i) An ***open cover*** for A is a family $(U_i)_{i\in I}$ of open subsets of V having the property that $A \subseteq \cup_{i\in I} U_i$.

(ii) A ***subcover*** of an open cover $(U_i)_{i\in I}$ of A is an open cover $(V_j)_{j\in J}$ of A having the property that $(V_j)_{j\in J} \subseteq (U_i)_{i\in I}$. •

We may now define compactness and other related properties of a subset of a normed vector space.

6.6.14 Definition (Bounded, compact, totally bounded) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. A subset $A \subseteq \mathsf{V}$ is:

(i) ***bounded*** if there exists $M \in \mathbb{R}_{>0}$ such that $A \subseteq \overline{\mathsf{B}}(M, 0)$;

(ii) ***compact*** if every open cover $(U_i)_{i\in I}$ of A possesses a finite subcover;

(iii) ***precompact***[3] if $\mathrm{cl}(A)$ is compact;

(iv) ***totally bounded*** if, for every $\epsilon \in \mathbb{R}_{>0}$ there exists $v_1, \dots, v_k \in \mathsf{V}$ such that $A \subseteq \cup_{j=1}^{k} \mathsf{B}(\epsilon, v_j)$. •

[3] What we call "precompact" is very often called "relatively compact." However, we shall use the term "relatively compact" for something different.

6.6.15 Theorem (Compactness and dimension) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. Then the following statements are equivalent:*

(i) V is finite-dimensional;

(ii) the closed unit ball $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is compact;

(iii) a subset $K \subseteq \mathsf{V}$ is compact if and only if it is closed and bounded;

(iv) V with the norm topology is locally compact.

Proof (i) $\Longrightarrow$ (ii) By Proposition 6.6.2 $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is closed and bounded. Now, if $\{e_1, \dots, e_n\}$ is a basis for V, we have a map $\iota\colon \mathsf{V} \to \mathbb{F}^n$ defined by

$$\iota(v_1e_1 + \cdots + v_ne_n) = (v_1, \dots, e_n)$$

which induces a norm on $\mathbb{F}^n$ (cf. the proof of Proposition 6.1.4), which we also denote by $\|\cdot\|$. Since ι is a homeomorphism of the normed vector spaces $(\mathsf{V}, \|\cdot\|)$ and $(\mathbb{F}^n, \|\cdot\|)$, it follows from the Heine–Borel Theorem that $\overline{\mathsf{B}}(1, 0_{\mathbb{F}^n})$ is compact. Since the image of compact sets under continuous maps is compact (Proposition **??**), we conclude that $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is compact.

(ii) $\Longrightarrow$ (iii) Suppose that $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is compact and let $K \subseteq \mathsf{V}$ be compact. By Proposition **??** we immediately have that K is closed. Let $\epsilon \in \mathbb{R}_{>0}$ and note that $(\mathsf{B}(\epsilon, v))_{v \in K}$ is an open cover of K. Then there exists a finite subset $v_1, \dots, v_k \in K$ such that

$$K \subseteq \mathsf{B}(\epsilon, v_1) \cup \cdots \cup \mathsf{B}(\epsilon, v_k).$$

We claim that $\cup_{j=1}^k \mathsf{B}(\epsilon, v_k)$ is bounded. Let

$$M = \max\{\|v_j\| \mid j, l \in \{1, \dots, k\}\} + \epsilon.$$

For $j \in \{1, \dots, k\}$ and $v \in \mathsf{B}(\epsilon, v_j)$ we compute

$$\|v\| = \|v - v_j + v_j\| \leq \|v - v_j\| + \|v_j\| \leq M.$$

Thus $\cup_{j=1}^k \mathsf{B}(\epsilon, v_k) \subseteq \overline{\mathsf{B}}(M, 0_{\mathsf{V}})$. Thus K is bounded as well as being closed.

Now suppose that $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is compact and let $K \subseteq \mathsf{V}$ be closed and bounded. Since K is bounded $K \subseteq \overline{\mathsf{B}}(M, 0_{\mathsf{V}})$ for some $M \in \mathbb{R}_{>0}$. By Proposition 6.6.2, $\overline{\mathsf{B}}(M, 0_{\mathsf{V}})$ is compact. Then K is a closed subset of a compact set, and so is compact by Proposition **??**.

(iii) $\Longrightarrow$ (iv) Since V is a metric space it is Hausdorff by Proposition **??**. Thus we need only show that $v \in \mathsf{V}$ possesses a precompact neighbourhood. However, for any $\epsilon \in \mathbb{R}_{>0}$, $\mathsf{B}(\epsilon, v)$ is a neighbourhood of v. We claim that $\overline{\mathsf{B}}(\epsilon, 0_{\mathsf{V}})$ is closed and bounded, and so compact by hypothesis. It is clearly bounded since $\overline{\mathsf{B}}(\epsilon, v) \subseteq \overline{\mathsf{B}}(M, 0_{\mathsf{V}})$ where $M = \|v\| + \epsilon$ (why?). It is moreover closed since, as we showed in the first part of the proof, it is the preimage of a closed set under a continuous map.

(iv) $\Longrightarrow$ (i) Let us first show that, if $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is compact, then V is finite-dimensional. Note that $(\mathsf{B}(\frac{1}{2}, v))_{v \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}})}$ is an open covering of $\overline{\mathsf{B}}(\frac{1}{2}, 0_{\mathsf{V}})$. Therefore, there exists $v_1, \dots, v_k \in \overline{\mathsf{B}}(\frac{1}{2}, 0_{\mathsf{V}})$ such that

$$\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \subseteq \mathsf{B}(\tfrac{1}{2}, v_1) \cup \cdots \cup \mathsf{B}(\tfrac{1}{2}, v_k).$$

Let $\mathsf{U} = \operatorname{span}_{\mathbb{R}}(v_1, \dots, v_k)$, which is then a finite-dimensional subspace of V. Since U is complete by Theorem 6.3.3 it is closed by Proposition **??**. We will show that $\mathsf{U} = \mathsf{V}$. Suppose this is not so and let $v_0 \in \mathsf{V} \setminus \mathsf{U}$. Since U is closed, $v_0 \notin \operatorname{cl}(\mathsf{U})$ and so by Proposition 6.6.8 the number

$$r = \inf\{\|u - v_0\| \mid u \in \mathsf{V}\}$$

is in $\mathbb{R}_{>0}$. Let $R \in \mathbb{R}_{>0}$ be such that $\overline{\mathsf{B}}(R, v_0) \cap \mathsf{U} \neq \emptyset$. Then $\overline{\mathsf{B}}(R, v_0) \cap \mathsf{U}$ is closed since it is the intersection of closed sets. The set $\overline{\mathsf{B}}(R, v_0) \cap \mathsf{U}$ is also clearly bounded. Since we have proved that (i) $\implies$ (iii) it follows that $\overline{\mathsf{B}}(R, v_0) \cap \mathsf{U}$ is compact. Define $f \colon \overline{\mathsf{B}}(R, v_0) \cap \mathsf{U} \to \mathbb{R}$ by $f(u) = \|u - v_0\|$. By Proposition 6.5.4 this function is continuous. By Theorem **??** it follows that f achieves its minimum on $\overline{\mathsf{B}}(R, v_0) \cap \mathsf{U}$. Since $R \geq r$ it follows that there exists $u_0 \in \overline{\mathsf{B}}(R, v_0) \cap \mathsf{U}$ such that $f(u_0 - v_0) = r$. Since $\frac{v_0 - u_0}{\|v_0 - u_0\|} \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ there is some $j \in \{1, \dots, k\}$ such that $\frac{v_0 - u_0}{\|v_0 - u_0\|} \in \mathsf{B}(\frac{1}{2}, v_j)$. Therefore,

$$\left\| \frac{v_0 - u_0}{\|v_0 - u_0\|} - v_j \right\| \leq \frac{1}{2} \quad \implies \quad \left\| v_0 - u_0 - \|v_0 - u_0\| v_j \right\| \leq \tfrac{1}{2}\|v_0 - u_0\| = \tfrac{r}{2}.$$

But we also have $v_0 - u_0 - \|v_0 - u_0\| v_j \in \mathsf{U}$ and so

$$\left\| v_0 - u_0 - \|v_0 - u_0\| v_j \right\| \geq r,$$

giving a contradiction. Thus $\mathsf{U} = \mathsf{V}$ and so compactness of $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ implies finite-dimensionality of V.

Now suppose that V is locally compact. Then there exists a neighbourhood U of 0_{V} for which $\operatorname{cl}(U)$ is compact. Openness of U implies that there exists $\epsilon \in \mathbb{R}_{>0}$ such that $\mathsf{B}(\epsilon, 0_{\mathsf{V}}) \subseteq U$. Then $\overline{\mathsf{B}}(\epsilon, 0_{\mathsf{V}})$ is a closed subset of the compact set $\operatorname{cl}(U)$, implying by Proposition **??** that $\overline{\mathsf{B}}(\epsilon, 0_{\mathsf{V}})$ is compact. By Proposition 6.6.2 it follows that $\overline{\mathsf{B}}(1, 0_{\mathsf{V}})$ is compact. Our argument above implies that V is finite-dimensional. ■

The theorem is rather an important one, given that compact sets have important properties that one often makes use of in applications (see, for example, ***missing stuff***). Since, in infinite dimensions, one loses the convenient interpretation of compact sets as being equivalent to closed and bounded sets, it then becomes important to understand the nature of the compact sets in a given normed vector space. This can really only be done on a case-by-case basis. For example, we do this in ***missing stuff*** for ***missing stuff***. The fact that the closed unit ball is not compact in infinite dimensions is also responsible for the sometimes nonintuitive distinctions between finite- and infinite-dimensional normed vector spaces. We shall try to point out specific instances of this as we go along.

6.6.4 Closed subspaces

Subspaces of normed vector spaces are again normed vector spaces by Proposition 6.1.7(iv). It is interesting to know what properties a subspace inherits from the space in which is sits. This is simple in finite-dimensions, but rather more complicated in infinite-dimensions. For reasons that are perhaps not *a priori* clear, closed subspaces play an important rôle in Banach space theory and practice, and for this reason we here study closed subspaces in a little detail.

First we characterise the closed subspaces of a Banach space in a manner completely analogous to Proposition **??** for metric spaces.

6.6.16 Proposition (Characterisations of closed subspaces) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. For a subspace $\mathsf{U} \subseteq \mathsf{V}$ and with $\|\cdot\|_{\mathsf{U}}$ the restriction of $\|\cdot\|$ to U, the following statements hold:*

(i) if V is a Banach space and if U is closed, then $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ is a Banach space;

(ii) if $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$ is a Banach space then U is closed.

Proof (i) If $(u_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in U, then this is also a Cauchy sequence V. Thus the sequence converges to some $v \in \mathsf{V}$. By Proposition 6.6.8(ii) it follows that $v \in \mathrm{cl}(\mathsf{U}) = \mathsf{U}$, and so U is complete.

(ii) Let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in U converging to $v \in \mathsf{V}$. This is a Cauchy sequence in V and so is also a Cauchy sequence in U, by definition of $\|\cdot\|_{\mathsf{U}}$. Therefore, $v \in \mathsf{U}$ since U is complete. By Proposition 6.6.8(ii) it follows that U is closed. ■

The result has the following useful corollaries. The first is simply a useful rewording of Proposition 6.6.16. But the result is nice, because it says that closed subspaces of Banach spaces are Banach spaces, and so closed subspaces are the proper notion of "subobject" when dealing with Banach spaces.

6.6.17 Corollary (Subspaces of Banach spaces are closed if and only if they are complete) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $(\mathsf{V}, \|\cdot\|)$ is a $\mathbb{F}$-Banach space, a subspace $\mathsf{U} \subseteq \mathsf{V}$ is closed if and only if it is complete.*

The next corollary provides some insight into how one should view the completion of a normed vector space.

6.6.18 Corollary (The closure of a subspace is a completion of the subspace) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. For a subspace $\mathsf{U} \subseteq \mathsf{V}$ denote by $\|\cdot\|_{\mathsf{U}}$ and $\|\cdot\|_{\mathrm{cl}(\mathsf{U})}$ the restriction of $\|\cdot\|$ to U and $\mathrm{cl}(\mathsf{U})$, respectively. Then $(\mathrm{cl}(\mathsf{U}), \|\cdot\|_{\mathrm{cl}(\mathsf{U})})$ is a completion of $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$.*

Proof It is clear that the inclusion map of U into $\mathrm{cl}(\mathsf{U})$ preserves the norm, i.e., that $\|u\|_{\mathsf{U}} = \|u\|_{\mathrm{cl}(\mathsf{U})}$. Moreover, by Proposition 6.6.8(ii) it follows that, given $v \in \mathrm{cl}(\mathsf{U})$, there exists a sequence $(u_j)_{j\in\mathbb{Z}_{>0}}$ converging to v. Thus $(\mathrm{cl}(\mathsf{U}), \|\cdot\|_{\mathrm{cl}(\mathsf{U})})$ is indeed a completion of $(\mathsf{U}, \|\cdot\|_{\mathsf{U}})$. ■

The next two corollaries concern finite-dimensional cases.

6.6.19 Corollary (Finite-dimensional subspaces are closed) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. If $\mathsf{U} \subseteq \mathsf{V}$ is a finite-dimensional subspace then U is closed.*

Proof By Theorem 6.3.3, U is complete, and so is closed by part (ii) of Proposition 6.6.16. ■

6.6.20 Corollary (Subspaces of finite-dimensional normed vector spaces are closed) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a finite-dimensional normed* $\mathbb{F}$*-vector space. If* $\mathsf{U} \subseteq \mathsf{V}$ *is subspace then* U *is closed.*

Proof Subspaces of finite-dimensional vector spaces are finite-dimensional, and so closed by Corollary 6.6.19. ■

Let us record the topological properties of the basic subspace operations of sum and intersection. For intersections the story is fairly simple.

6.6.21 Proposition (Intersections of closed subspaces) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *if* $(\mathsf{V}, \|\cdot\|)$ *is a normed* $\mathbb{F}$*-vector space, and if* $(\mathsf{U}_{\mathrm{a}})_{\mathrm{a}\in\mathrm{A}}$ *is a family of closed subspaces of* V, *then* $\cap_{\mathrm{a}\in\mathrm{A}}$ *is a closed subspace of* V.

Proof The set $\cap_{a\in A}\mathsf{U}_a$ is a subspace by Proposition 4.3.34 and is closed by Proposition **??**. ■

For sums the story is significantly more complex. First we give a counterexample to the simplest statement one may wish to make.

6.6.22 Example (The sum of closed subspaces may not be closed) The example we use here begins with the Banach space $\ell^2(\mathbb{F})$ consisting of sequences $(a_j)_{j\in\mathbb{Z}_{>0}}$ for which $\sum_{j=1}^{\infty}|a_j|^2 < \infty$. The norm we use is

$$\|(a_j)_{j\in\mathbb{Z}_{>0}}\|_2 = \Big(\sum_{j=1}^{\infty}|a_j|^2\Big)^{1/2}.$$

In Corollary 6.7.21 we show that this is a Banach space and is, moreover, the completion of $\mathbb{F}_0^\infty$ under the norm $\|\cdot\|_2$. We denote by $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ the standard basis for $\mathbb{F}_0^\infty$. Thus

$$\boldsymbol{e}_j(k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

For the purposes of this example we consider two subspaces of $\ell^2(\mathbb{F})$. We let

$$\begin{aligned} \mathsf{U} &= \operatorname{cl}(\operatorname{span}_{\mathbb{F}}(\boldsymbol{e}_{2j-1}|\ j \in \mathbb{Z}_{>0})), \\ \mathsf{V} &= \operatorname{cl}(\operatorname{span}_{\mathbb{F}}(\boldsymbol{e}_{2j-1} + \tfrac{1}{j}\boldsymbol{e}_{2j}|\ j \in \mathbb{Z}_{>0})) \end{aligned}$$

so that U and V are closed subspaces.

Let us establish some facts about these subspaces via a sequence of lemmata.

1 Lemma $\mathsf{U} \cap \mathsf{V} = \{0_{\ell^2(\mathbb{F})}\}$.

Proof Let us denote

$$\mathsf{U}' = \operatorname{span}_{\mathbb{F}}(\boldsymbol{e}_{2j-1}|\ j \in \mathbb{Z}_{>0}), \quad \mathsf{V}' = \operatorname{span}_{\mathbb{F}}(\boldsymbol{e}_{2j-1} + \tfrac{1}{j}\boldsymbol{e}_{2j}|\ j \in \mathbb{Z}_{>0}).$$

Let $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{U} \cap \mathsf{V}$. By definition of U and V there exist sequences $((x_{jl})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ and $((y_{jl})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ in U' and V', respectively, such that

$$\lim_{l\to\infty}(x_{jl})_{j\in\mathbb{Z}_{>0}} = \lim_{l\to\infty}(y_{jl})_{j\in\mathbb{Z}_{>0}} = (a_j)_{j\in\mathbb{Z}_{>0}}.$$

Since $(x_{jl})_{j\in\mathbb{Z}_{>0}} \in \mathsf{U}'$ for each $l \in \mathbb{Z}_{>0}$ it follows that $a_{2j} = 0$ for $j \in \mathbb{Z}_{>0}$. Therefore, $\lim_{l\to\infty} y_{(2j)l} = 0$ for $j \in \mathbb{Z}_{>0}$. Since $y_{(2j-1)l} = jy_{(2j)l}$ for each $j, l \in \mathbb{Z}_{>0}$ it then follows that $\lim_{l\to\infty} y_{(2j-1)l} = 0$ for each $j \in \mathbb{Z}_{>0}$. Therefore, $a_j = 0$ for each $j \in \mathbb{Z}_{>0}$, giving the lemma. ▼

2 Lemma $\mathrm{cl}(\mathsf{U} + \mathsf{V}) = \ell^2(\mathbb{F})$.

Proof Let U' and V' be as in the proof of Lemma 1. We claim that $\mathsf{U}' + \mathsf{V}' = \mathbb{F}_0^\infty$. To see this, let $(a_j)_{j\in\mathbb{Z}_{>0}}$ and write $a_j = x_j + y_j$ where

$$x_j = \begin{cases} a_j - ja_{j+1}, & j \text{ odd}, \\ 0, & j \text{ even}, \end{cases} \qquad y_j = \begin{cases} ja_{j+1}, & j \text{ odd}, \\ a_j, & j \text{ even}. \end{cases}$$

Note that $(x_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{U}'$ and $(y_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{V}'$. Thus $\mathsf{U}' + \mathsf{V}' = \mathbb{F}_0^\infty$, as desired. Therefore,

$$\mathrm{cl}(\mathsf{U}' + \mathsf{V}') = \mathrm{cl}(\mathbb{F}_0^\infty) = \ell^2(\mathbb{F})$$

and so

$$\mathrm{cl}(\mathsf{U}' + \mathsf{V}') \subseteq \mathrm{cl}(\mathsf{U} + \mathsf{V}) = \ell^2(\mathbb{F}),$$

as desired. ▼

3 Lemma $\mathsf{U} + \mathsf{V} \subset \ell^2(\mathbb{F})$.

Proof Following the proof of Lemma 1, elements of U and V have the form

$$(x_1, 0, x_2, 0, x_3, 0, \dots), \qquad (y_1, y_1, y_2, \tfrac{1}{2}y_2, y_3, \tfrac{1}{3}y_3, \dots),$$

respectively, where

$$\sum_{j=1}^\infty |x_j|^2 < \infty, \quad \sum_{j=1}^\infty |y_j|^2 + \sum_{j=1}^\infty \frac{|y_j|^2}{j^2} < \infty. \tag{6.7}$$

Thus an element of $\mathsf{U} + \mathsf{V}$ has the form

$$(x_1 + y_1, y_1, x_2 + y_2, \tfrac{1}{2}y_2, x_3 + y_3, \tfrac{1}{3}y_3, \dots), \tag{6.8}$$

where the inequalities (6.7) hold. Now consider the sequence

$$(1, 1, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{3}, \dots) \in \ell^2(\mathbb{F}).$$

We claim that this sequence is not in $\mathsf{U} + \mathsf{V}$. Indeed, suppose that the sequence can be expressed in the form (6.8). Then we must have $x_j + y_j = \frac{1}{j}$ and $\frac{1}{j}y_j = \frac{1}{j}$ for each $j \in \mathbb{Z}_{>0}$. Thus $x_j = \frac{1}{j} - 1$ and $y_j = 1$. The inequalities (6.7) do not hold in this case, so the sequence cannot be in $\mathsf{U} + \mathsf{V}$. ▼

Now we make the following observation. The subspaces U and V are closed and complementary. The sum $\mathsf{U} + \mathsf{V}$ is a strict subspace of $\ell^2(\mathbb{F})$ but is dense in $\ell^2(\mathbb{F})$. Thus $\mathsf{U} + \mathsf{V} \subset \mathrm{cl}(\mathsf{U} + \mathsf{V})$ and so $\mathsf{U} + \mathsf{V}$ is not closed. That is, the sum of closed subspaces need not be closed. •

Now being deprived of access to the nicest result concerning sums of closed subspaces, we must now wonder what *is* true. It turns out that the story here is a little complicated, but it is worth understanding since it actually reveals something interesting about Banach space geometry. So let us spend a few moments understanding this. Suppose that we have a Banach space $(\mathsf{V}, \|\cdot\|)$ with two closed subspaces U_1 and U_2. Then define

$$\delta(\mathsf{U}_1, \mathsf{U}_2) = \sup\{\rho \in [0,1] \mid \overline{\mathsf{B}}(\rho, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2) \subseteq (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2\},$$

with the convention that if $A, B \subseteq \mathsf{V}$ then

$$A + B = \{u + v \mid u \in A, \ v \in B\}.$$

This is a definition with geometric character so let us examine it in a simple case so that we have a little insight into what it means.

6.6.23 Example ($\delta(\mathsf{U}_1, \mathsf{U}_2)$) Let $\mathsf{V} = \mathbb{R}^2$ and let

$$\mathsf{U}_1 = \operatorname{span}_{\mathbb{R}}((1,0)), \quad \mathsf{U}_2 = \operatorname{span}_{\mathbb{R}}((1,1)).$$

We use the standard norm on $\mathbb{R}^2$: $\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}$. The set $(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$ is depicted on the left in Figure 6.2 and $\overline{\mathsf{B}}(\rho, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2)$ is shown on the right.

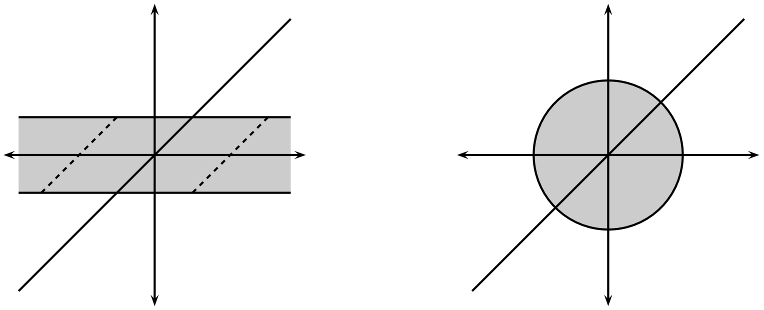


Figure 6.2 The definition of $\delta(\mathsf{U}_1, \mathsf{U}_2)$: $(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$ on the left and $\overline{\mathsf{B}}(\rho, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2)$ on the right

The idea is that $(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$ is obtained by translating the unit ball in U_1 by all vectors in U_2. Thus one "thickens" U_2 by the unit ball in U_1. Now one take balls of increasing radius in $\mathsf{U}_1 + \mathsf{U}_2$ until the ball is no longer contained in $(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$. In this example one can see that $\delta(\mathsf{U}_1, \mathsf{U}_2) = 1$. •

In finite dimensions the constructions we give are not so insightful. For example, if V is finite-dimensional then $\delta(\mathsf{U}_1, \mathsf{U}_2) > 0$. However, in infinite dimensions it turns out that $\delta(\mathsf{U}_1, \mathsf{U}_2)$ measures when $\mathsf{U}_1 + \mathsf{U}_2$ is not closed.

6.6.24 Theorem (When is the sum of closed subspaces closed?) *If* $(\mathsf{V}, \|\cdot\|)$ *is a Banach space and if* U_1 *and* U_2 *are closed subspaces of* V*, then* $\mathsf{U}_1 + \mathsf{U}_2$ *is closed if and only if* $\delta(\mathsf{U}_1, \mathsf{U}_2) > 0$.

Proof Let us define

$$\begin{aligned}\alpha(\mathsf{U}_1, \mathsf{U}_2) =\ & \sup\{\rho \in [0,1] \mid \overline{\mathsf{B}}(\rho, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq \mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2)\},\\ \beta(\mathsf{U}_1, \mathsf{U}_2) =\ & \sup\{\rho \in [0,1] \mid \overline{\mathsf{B}}(\rho, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2\}.\end{aligned}$$

Both of these quantities are, in fact, equal to $\delta(\mathsf{U}_1, \mathsf{U}_2)$.

1 **Lemma** $\alpha(\mathsf{U}_1, \mathsf{U}_2) = \beta(\mathsf{U}_1, \mathsf{U}_2) = \delta(\mathsf{U}_1, \mathsf{U}_2)$.

Proof Let us abbreviate

$$\alpha = \alpha(\mathsf{U}_1, \mathsf{U}_2), \quad \beta = \beta(\mathsf{U}_1, \mathsf{U}_2), \quad \delta = \delta(\mathsf{U}_1, \mathsf{U}_2).$$

Let us first prove that $\alpha \leq \beta$. This is clearly true if $\alpha = 0$ so suppose that $\alpha > 0$. Since $\mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2)$ is closed we have

$$\overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq \mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2).$$

Note that

$$(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2 = \cap_{r \in (0,1)}((\overline{\mathsf{B}}(\tfrac{1}{1-r}, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2).$$

Therefore, if

$$\overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq (\overline{\mathsf{B}}(\tfrac{1}{1-r}, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2 \tag{6.9}$$

for every $r \in (0,1)$ then we have

$$\overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2 \tag{6.10}$$

since $\overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2)$ is closed. Moreover, if (6.9) holds then $\alpha \leq \beta$, and so it thus suffices for this part of the proof to show that (6.9) holds for every $r \in (0,1)$. By Proposition 6.6.8 we have

$$\mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2) \subseteq ((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2) + \overline{\mathsf{B}}(\alpha r, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2)$$

for every $r \in (0,1)$. By definition of α we then have

$$\begin{aligned}\overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) &\subseteq \mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2)\\ &\subseteq ((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2) + \overline{\mathsf{B}}(\alpha r, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \end{aligned} \tag{6.11}$$

for every $r \in (0,1)$. Let

$$u_0 \in \overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2).$$

By (6.11) there exists

$$u_1 \in \overline{\mathsf{B}}(\alpha r, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2), \quad v_0 \in (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$$

such that $u_0 = v_0 + u_1$. By definition of α we have

$$\overline{\mathsf{B}}(\alpha r, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq r\,\mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2),$$

where, if $a \in \mathbb{F}$ and $A \subseteq \mathsf{V}$, we denote $aA = \{av \mid v \in A\}$. Thus, again by (6.11), there exists

$$u_2 \in \overline{\mathsf{B}}(\alpha r^2, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2), \quad v_1 \in (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$$

such that $u_1 = rv_1 + u_2$. Continuing in this way, there exist sequences $(u_j)_{j\in\mathbb{Z}_{\geq 0}}$ and $(v_j)_{j\in\mathbb{Z}_{\geq 0}}$ such that

1. $u_j \in \overline{\mathsf{B}}(\alpha r^j, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq r^j \,\mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2)$,
2. $v_j \in (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2$, and
3. $u_j = r^j v_j + v_{j+1}$

for each $j \in \mathbb{Z}_{>0}$. Clearly, then, $\lim_{j\to\infty} u_j = 0_{\mathsf{V}}$. Therefore,

$$u_0 - u_k = \sum_{j=0}^{k} q^j v_j \quad \Longrightarrow \quad u_0 = \lim_{k\to\infty}(u_0 - u_k) = \sum_{j=0}^{\infty} r^j v_j.$$

Also,

$$\|v_j\| = r^{-j}\|u_j - u_{j+1}\| \leq r^{-j}(\|u_j\| + u_{j+1}) \leq r^{-j}(\alpha r^j + \alpha r^{j+1}) = \alpha(1 + r).$$

Thus the sequence $(v_j)_{j\in\mathbb{Z}_{\geq 0}}$ is bounded. Now, for each $j \in \mathbb{Z}_{\geq 0}$ define $w_j \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1$ and $z_j \in \mathsf{U}_2$ such that $v_j = w_j + z_j$. Then we have

$$\|z_j\| = \|v_j - w_j\| \leq \|v_j\| + \|w_j\| \leq \alpha(1 + r) + 1,$$

and so the sequence $(z_j)_{j\in\mathbb{Z}_{\geq 0}}$ is bounded. Therefore,

$$\Big\|\sum_{j=0}^{\infty} r^j w_j\Big\| \leq \sum_{j=0}^{\infty} r^j\|w_j\| \leq \frac{1}{1 - r} \quad \Longrightarrow \quad \sum_{j=0}^{\infty} r^j w_j \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1$$

since $\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1$ is closed. Similarly,

$$\Big\|\sum_{j=0}^{\infty} r^j z_j\Big\| \leq \sum_{j=0}^{\infty} r^j\|z_j\| \leq \frac{1}{1 - r}(\alpha(1 + r) + 1) \quad \Longrightarrow \quad \sum_{j=0}^{\infty} r^j z_j \in \mathsf{U}_2$$

since U_2 is closed. Thus

$$u_0 = \sum_{j=0}^{\infty} r^j v_j = \sum_{j=0}^{\infty} r^j w_j + \sum_{j=0}^{\infty} r^j z_j \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1 + \mathsf{U}_2.$$

Since u_0 was chosen arbitrarily from $\overline{\mathsf{B}}(\alpha, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2)$ and since the argument can be made for every $r \in (0, 1)$, we have shown that (6.9) holds for every $r \in (0, 1)$, giving $\alpha \leq \beta$.

That $\beta \leq \delta$ follows directly from the definitions.

To show that $\delta \leq \alpha$, and so to complete the proof, it suffices to show that

$$\mathrm{cl}(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2)) = \overline{\mathsf{B}}(1, 0_{\mathsf{V}})\,\mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2)$$

(why?). To show this, let $v \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}})\,\mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2)$. If $v = 0_{\mathsf{V}}$ then we obviously have $v \in \mathrm{cl}(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2))$. Thus we can suppose that $v \neq 0_{\mathsf{V}}$. Let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{U}_1 + \mathsf{U}_2$ converging to v. We can without loss of generality suppose that $v_j \neq 0_{\mathsf{V}}$ for each $j \in \mathbb{Z}_{>0}$. Then define $u_j = \frac{\|v\|}{\|v_j\|} v_j$ for each $j \in \mathbb{Z}_{>0}$, noting that $u_j \in \overline{\mathsf{B}}(1, 0_{\mathsf{V}})\,\mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2)$. Moreover, $\lim_{j\to\infty} u_j = v$ and so $v \in \mathrm{cl}(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2))$. This gives

$$\overline{\mathsf{B}}(1, 0_{\mathsf{V}})\,\mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq \mathrm{cl}(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2)).$$

By Proposition 6.6.10 we have

$$\mathrm{cl}(\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap (\mathsf{U}_1 + \mathsf{U}_2)) \subseteq \overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2).$$

This gives $\delta < \alpha$ be the definitions. ▼

Carrying on with the proof of the theorem, first suppose that $\mathsf{U}_1 + \mathsf{U}_2$ is closed. By Corollary 6.6.17 it follows that $\mathsf{U}_1 + \mathsf{U}_2$ is complete. We obviously have

$$\mathsf{U}_1 + \mathsf{U}_2 = \cup_{j=1}^{\infty} j((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2).$$

Therefore, by the Baire Category Theorem***missing stuff*** there exists at least one $j \in \mathbb{Z}_{>0}$ for which

$$\mathrm{int}(\mathrm{cl}(j((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2))) \neq \emptyset.$$

Thus there exist $v \in \mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2)$ and $r \in \mathbb{R}_{>0}$ such that

$$\overline{\mathsf{B}}(r, v) \cap (\mathsf{U}_1 + \mathsf{U}_2) \subseteq j\,\mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2).$$

Therefore,

$$\overline{\mathsf{B}}(\tfrac{r}{j}, v) \cap (\mathsf{U}_1 + \mathsf{U}_2) \subseteq \mathrm{cl}((\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2),$$

giving $\alpha(\mathsf{U}_1, \mathsf{U}_1) > 0$ and so, by the lemma, $\delta(\mathsf{U}_1, \mathsf{U}_2) > 0$.

Conversely, suppose that $\delta(\mathsf{U}_1, \mathsf{U}_2) > 0$ and so, by the lemma, $\beta(\mathsf{U}_1, \mathsf{U}_2) > 0$. Let $\beta \in (0, \beta(\mathsf{U}_1, \mathsf{U}_2))$. We obviously have

$$\mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) = \cup_{j=1}^{\infty} j(\overline{\mathsf{B}}(\beta, 0_{\mathsf{V}})) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2).$$

By definition of $\beta(\mathsf{U}_1, \mathsf{U}_2)$ it holds that

$$\overline{\mathsf{B}}(\beta, 0_{\mathsf{V}}) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq (\overline{\mathsf{B}}(1, 0_{\mathsf{V}}) \cap \mathsf{U}_1) + \mathsf{U}_2.$$

Moreover, we then obviously have

$$\cup_{j=1}^{\infty} j(\overline{\mathsf{B}}(\beta, 0_{\mathsf{V}})) \cap \mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq \mathsf{U}_1 + \mathsf{U}_2.$$

This gives $\mathrm{cl}(\mathsf{U}_1 + \mathsf{U}_2) \subseteq \mathsf{U}_1 + \mathsf{U}_2$ which gives $\mathsf{U}_1 + \mathsf{U}_2$ as being closed, as desired. ■

Let us close our discussion of closed subspaces by considering some examples of closed and non-closed subspaces of normed vector spaces.

6.6.25 Examples (Closed subspace) Both examples we consider are subspaces of the normed vector space $(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R}), \|\cdot\|_\infty)$. By $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ we denote the set of bounded, continuous $\mathbb{R}$-valued functions on $\mathbb{R}$ and the norm $\|\cdot\|_\infty$ is defined thusly:

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in \mathbb{R}\}.$$

Note that convergence in the norm $\|\cdot\|_\infty$ is, by definition, uniform convergence. In Theorem 3.5.8 we essentially showed that $(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R}), \|\cdot\|_\infty)$ is a Banach space, although we shall revisit this in Section 6.7.4.

1. Let $\mathsf{C}^0_0(\mathbb{R};\mathbb{R})$ be the subset of $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ consisting of those functions satisfying

$$\lim_{x\to-\infty} f(x) = 0, \quad \lim_{x\to\infty} f(x) = 0.$$

It is easy to verify (cf. Proposition 2.3.23) that $\mathsf{C}^0_0(\mathbb{R};\mathbb{R})$ is a subspace of $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$. We claim that it is a closed subspace. To show this, it suffices to show that, if $(f_j)_{j\in\mathbb{Z}_{>0}}$ is any sequence in $\mathsf{C}^0_0(\mathbb{R};\mathbb{R})$ converging in $\mathsf{C}_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$, then the limit function is in $\mathsf{C}^0_0(\mathbb{R};\mathbb{R})$. We shall prove this below as Theorem 6.7.40, but it is not too hard to imagine why it is true. Uniform convergence requires that the limit function be approximated uniformly over all of $\mathbb{R}$ by sufficiently large terms in the sequence. Since all functions in the sequence tend to zero at infinity, they will pull the limit function down to zero with them.

2. Let $\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ be the subset of $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ consisting of those functions that are continuously differentiable. By Proposition 3.2.10 it follows that $\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ is a subspace of $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$. We claim that it is not closed. To see this, define a sequence of functions $(f_j)_{j\in\mathbb{Z}_{>0}}$ as follows:

$$f_j(x) = \begin{cases} -1, & x \in (-\infty, -1-\frac{1}{j}), \\ \frac{1}{4}jx^2 + \frac{1}{2}(j+1)x + \frac{(j-1)^2}{4j}, & x \in [-1-\frac{1}{j}, -1+\frac{1}{j}], \\ x, & x \in (-1+\frac{1}{j}, 1-\frac{1}{j}), \\ -\frac{1}{4}jx^2 + \frac{1}{2}(j+1)x - \frac{(j-1)^2}{4j}, & x \in [1-\frac{1}{j}, 1+\frac{1}{j}], \\ 1, & x \in (1+\frac{1}{j}, \infty). \end{cases}$$

We depict this sequence in Figure 6.3. One can show by direct computation that f_j is differentiable for each $j \in \mathbb{Z}_{>0}$; one need only check that the left and right limits for the function and its derivative match at the points $-1-\frac{1}{j}$, $-1+\frac{1}{j}$, $1-\frac{1}{j}$, and $1+\frac{1}{j}$. A direct computation also shows that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges pointwise to the function

$$f(x) = \begin{cases} -1, & x \in (-\infty, -1), \\ x, & x \in [-1, 1], \\ 1, & x \in (1, \infty). \end{cases}$$

To show that $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to f in $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ we need to show that the convergence is uniform.

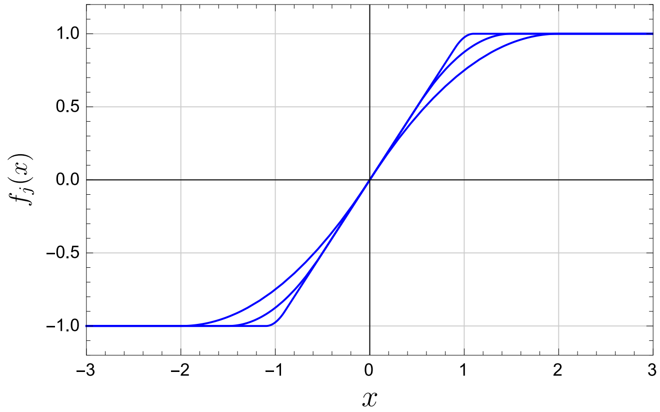


Figure 6.3 A sequence in $\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ not converging in $\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ (the terms f_1, f_2, and f_{10} are shown)

This is sort of "obvious" from Figure 6.3, but let us go through the details anyway. The only possible problems can occur on the intervals $[-1-\frac{1}{j},-1+\frac{1}{j}]$ and $[1-\frac{1}{j},1+\frac{1}{j}]$ since off these intervals f_j agrees with f. So let $\epsilon\in\mathbb{R}_{>0}$ and let N be sufficiently large that $\frac{1}{4N}<\epsilon$. On $[1-\frac{1}{j},1]$ the maximum deviation of f_j from f will occur at $x=1$. Thus, for $x\in[1-\frac{1}{j},1]$ we have

$$\begin{aligned}|f_j(x)-f(x)| &\le |f_j(1)-f(1)|\\ &= \left|-\tfrac{1}{4}j+\tfrac{1}{2}(j+1)-\tfrac{(j-1)^2}{4j}-1\right| = \left|\tfrac{1}{4j}\right| < \epsilon\end{aligned}$$

for $j\ge N$. Similarly, on $[1,1+\frac{1}{j}]$ the maximum deviation of f_j from f will occur at $x=1$, and the same computation gives $|f_j(x)-f(x)|<\epsilon$ for $x\in[1,1+\frac{1}{j}]$ for $j\ge N$. This gives $|f_j(x)-f(x)|<\epsilon$ for $x\in[1-\frac{1}{j},1+\frac{1}{j}]$. An entirely similar argument gives $|f_j(x)-f(x)|<\epsilon$ for $x\in[-1-\frac{1}{j},-1+\frac{1}{j}]$.

The point is that the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ converges to $f\in\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$. But since $f\notin\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$, it follows that $\mathsf{C}^1_{\mathrm{bdd}}(\mathbb{R};\mathbb{R})$ is not closed. The reason for this is fairly evident. The norm $\|\cdot\|_\infty$ does not know anything about the derivative of a function, and so it cannot be expected that the sequence of derivatives will converge to the derivative of the limit function, nor even that the limit function will indeed be even differentiable. •

6.6.5 Bases for normed vector spaces

In Section 4.3.4 we discussed at length the notion of a basis for a vector space, sometimes called a Hamel basis. The fact that every vector space possesses a Hamel basis is of great use in algebra, but not great value in analysis. To exhibit the limitations of the effectiveness of Hamel bases, let us prove that certain vector spaces are incapable of supporting a norm for which the resulting normed vector space is complete (thus we are supposing here familiarity with completeness, a notion we discuss in detail in Section 6.3).

6.6.26 Theorem (Dimension of an infinite-dimensional Banach space) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is an infinite-dimensional* $\mathbb{F}$*-Banach space, then* $\dim_{\mathbb{F}}(\mathsf{V}) \geq \mathrm{card}(\mathbb{R})$*. If* V *is separable then* $\dim_{\mathbb{F}}(\mathsf{V}) = \mathrm{card}(\mathbb{R})$*.*

Proof Let $v_1 \in \mathsf{V} \setminus \{0_{\mathsf{V}}\}$ be such that $\|v_1\| = 1$. Define $\hat{\alpha}_1 \colon \mathrm{span}_{\mathbb{F}}(v_1) \to \mathbb{F}$ by $\hat{\alpha}_1(av_1) = a$. It is trivial to check that $\hat{\alpha}_1$ is a continuous linear function satisfying $\hat{\alpha}_1(v_1) = 1$. By the Hahn–Banach Theorem, Theorem **??**, there exists $\alpha_1 \in \mathsf{V}^*$ such that $\alpha_1(v_1) = 1$. Next consider the closed subspace $\mathsf{V}_2 = \ker(\alpha_1)$ and let $v_2 \in \mathsf{V}_2$ so that $\alpha_1(v_2) = 0$. Also suppose that $\|v_2\| = 1$. Then define $\hat{\alpha}_2 \colon \mathrm{span}_{\mathbb{F}}(v_1, v_2) \to \mathbb{F}$ by $\hat{\alpha}_2(a_1v_1 + a_2v_2) = a_2$. As above, use the Hahn–Banach Theorem to deduce the existence of $\alpha_2 \in \mathsf{V}^*$ such that $\alpha_2(a_1v_2 + a_2v_2) = a_2$ for every $a_1, a_2 \in \mathbb{F}$. We may continue inductively in this way to define sequences $(v_j)_{j\in\mathbb{Z}_{>0}}$ and $(\alpha_j)_{j\in\mathbb{Z}_{>0}}$ such that $\|v_j\| = 1$, $j \in \mathbb{Z}_{>0}$, and such that

$$\alpha_j(v_k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

We claim that the family $(v_j)_{j\in\mathbb{Z}_{>0}}$ is linearly independent. Indeed, suppose that

$$c_1v_{j_1} + \cdots + c_kv_{j_k} = 0$$

for some $c_1, \dots, c_k \in \mathbb{F}$ and $j_1, \dots, j_k \in \mathbb{Z}_{>0}$. For each $l \in \{1, \dots, k\}$, apply α_{j_l} to the preceding equality to get $c_l = 0$. This give the desired linear independence. We also claim that

$$v_k \notin \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(v_j|\ \ j \neq k))$$

for each $k \in \mathbb{Z}_{>0}$. Indeed, if $(w_l)_{l\in\mathbb{Z}_{>0}}$ is a convergent sequence in $\mathrm{span}_{\mathbb{F}}(v_j|\ \ j \neq k)$ then $\alpha_k(w_l) = 0$ for all $l \in \mathbb{Z}_{>0}$. Continuity of α_k and Theorem 6.5.2 ensure that

$$\alpha_k(\lim_{l\to\infty} w_l) = \lim_{l\to\infty} \alpha_k(w_l) = 0.$$

Thus $\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(v_j|\ \ j \neq k)) \subseteq \ker(\alpha_k)$. Since $\alpha_k(v_k) = 1$ our claim follows.

Now we use a lemma.

1 Lemma *If* S *is a countably infinite set then there exists a family* $(\mathrm{A_t})_{\mathrm{t}\in[0,1]}$ *of infinite subsets of* S *such that* $\mathrm{A_{t_1}} \cap \mathrm{A_{t_2}}$ *is finite for* $\mathrm{t_1} \neq \mathrm{t_2}$*.*

Proof For $\theta \in [0, \pi)$ denote

$$\Sigma_\theta = \{(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \in \mathbb{R}^2 \mid x \in \mathbb{R},\ y \in [-1, 1]\}.$$

Thus Σ_θ is a bi-infinite strip of width 2 inclined at an angle θ to the x-axis in $\mathbb{R}^2$. For $\theta \in [0, \pi)$ define

$$\hat{A}_\theta = \{(x, y) \in \mathbb{Z}^2 \subseteq \mathbb{R}^2 \mid (x, y) \in \Sigma_\theta\}$$

as the points in $\mathbb{Z}^2$ lying in Σ_θ. Some elementary geometry can be used to verify the fact that if $\theta_1 \neq \theta_2$ then $\Sigma_{\theta_1} \cap \Sigma_{\theta_2}$ is compact. From this fact it follows that $\hat{A}_{\theta_1} \cap \hat{A}_{\theta_2}$ is finite for $\theta_1 \neq \theta_2$. Moreover, one can verify that $\hat{A}_\theta$ is infinite for every θ. To see this note that every ball of the form $\overline{\mathsf{B}}(1, (r\cos\theta, r\sin\theta))$ must contain a point with integer coordinates.

Since S and $\mathbb{Z}^2$ are both countable there exists a bijection $\phi\colon S \to \mathbb{Z}^2$. Since $[0,1]$ and $[0,\pi)$ both have the cardinality of $\mathbb{R}$ (why?), there exists a bijection $\psi\colon [0,1] \to [0,\pi)$. Then, for $t \in [0,1]$, define

$$A_t = \{s \in S \mid \phi(s) \in \hat{A}_{\psi(t)}\}.$$

It then follows that A_t is infinite since $\hat{A}_{\psi(t)}$ is infinite and that $A_{t_1} \cap A_{t_2}$ is finite since $\hat{A}_{\psi(t_1)} \cap \hat{A}_{\psi(t_2)}$ is finite. ▼

Now, using the lemma, let $(A_t)_{t\in[0,1]}$ be a family of subsets of $\mathbb{Z}_{>0}$ such that $A_{t_1} \cap A_{t_2}$ is finite for $t_1 \neq t_2$. Then define

$$u_t = \sum_{j\in A_t} \frac{v_j}{2^j}, \qquad t \in [0,1].$$

Note that

$$\|u_t\| = \Big\|\sum_{j\in A_t} \frac{v_j}{2^j}\Big\| \leq \sum_{j\in A_t} \frac{\|v_j\|}{2^j} < \infty$$

by Example 2.4.2–**??**. Thus the series for u_t is absolutely convergent and so convergent by Theorem 6.4.6. We claim that the set $\{u_t\}_{t\in[0,1]}$ is linearly independent. For $l \in \{1,\dots,k\}$ and $m \in \mathbb{Z}_{>0}$ we have

$$\alpha_m(u_{t_l}) = \alpha_m\Big(\sum_{j\in A_t} \frac{v_j}{2^j}\Big) = \sum_{j\in A_t} \frac{\alpha_m(v_j)}{2^j},$$

using Theorem 6.5.2. Thus

$$\alpha_m(u_{t_l}) = \begin{cases} 2^{-m}, & m \in A_{t_l}, \\ 0, & m \notin A_{t_l}. \end{cases}$$

Now suppose that

$$c_1 u_{t_1} + \dots + c_k u_{t_k} = 0 \tag{6.12}$$

for $c_1,\dots,c_k \in \mathbb{F}$ and $t_1,\dots,t_k \in [0,1]$. Without loss of generality we may suppose that the numbers $t_1,\dots,t_k$ are distinct. Then $\cap_{l=1}^k A_{t_l}$ is finite; let us denote it by $\{m_1,\dots,m_r\}$. For $l \in \{1,\dots,k\}$ define $A'_l = A_{t_l} \setminus \{m_1,\dots,m_r\}$, noting that the sets A'_l, $l \in \{1,\dots,k\}$, are countably infinite and disjoint. We can then rewrite (6.12) as

$$a_1 v_{m_1} + \dots + a_r v_{m_r} + c_1 \sum_{j_1\in A'_1} \frac{v_{j_1}}{2^{j_1}} + \dots + c_k \sum_{j_k\in A'_k} \frac{v_{j_k}}{2^{j_k}}$$

for suitable constants $a_1,\dots,a_r \in \mathbb{F}$ that depend on the coefficients $c_1,\dots,c_k$ and factors of $\frac{1}{2}$; the precise form of these is immaterial to our computations. Indeed, for each $l \in \{1,\dots,k\}$ let $m_l \in A'_l$. Then, by the properties for $(v_j)_{j\in\mathbb{Z}_{>0}}$ and $(\alpha_j)_{j\in\mathbb{Z}_{>0}}$ given before the lemma,

$$0 = \alpha_{m_l}(c_1 u_{t_1} + \dots + c_k u_{t_k}) = \frac{c_l}{2^{m_l}}.$$

Thus $c_l = 0$ for each $l \in \{1,\dots,k\}$, giving linear independence of $\{u_t\}_{t\in[0,1]}$. Since $\operatorname{card}([0,1]) = \operatorname{card}(\mathbb{R})$ the first assertion of the theorem follows.

For the final assertion of the theorem we shall prove that $\operatorname{card}(\mathsf{V}) = \operatorname{card}(\mathbb{R})$ if V is separable. It is clear that $\operatorname{card}(\mathsf{V}) \geq \operatorname{card}(\mathbb{R})$. For the opposite inequality, let $D \subseteq \mathsf{V}$ be a countable dense subset of V. For $v \in \mathsf{V}$ we can write $v = \lim_{j\to\infty} v_j$ for a sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ in D. Thus to every point in V we assign a sequence in the countable set D. The set of such sequences is $D^{\mathbb{Z}_{>0}}$, and so $\operatorname{card}(\mathsf{V}) \leq \operatorname{card}(D^{\mathbb{Z}_{>0}}) = \aleph_0^{\aleph_0}$. Now note that $2 \leq \aleph_0 \leq 2^{\aleph_0}$ by Example **??–??** and Exercise **??**. Thus

$$2^{\aleph_0} \leq \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0\cdot\aleph_0} = 2^{\aleph_0}$$

by Theorem **??**. Thus $\aleph_0^{\aleph_0} = 2^{\aleph_0}$ and so $\operatorname{card}(\mathsf{V}) \leq 2^{\aleph_0} = \operatorname{card}(\mathbb{R})$ by Exercise **??**. ■

6.6.27 Corollary (There are no Banach spaces of countable dimension) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if V is an $\mathbb{F}$-vector space with an infinite, countable Hamel basis, then there is no norm on V for which the resulting normed $\mathbb{F}$-vector space is complete.*

missing stuff

6.6.6 Notes

Our approach to characterising the closedness of sums of closed subspaces follows **RM/BS:79**, who base their presentation on that of **TK:80a**. Note that we also used this characterisation of sums of closed subspaces in our proofs of the Open Mapping Theorem and the Closed Graph Theorem. This idea is included in the paper of **RM/BS:79**.

The proof we give for Theorem 6.6.26 is due to **HEL:73**. The proof of the lemma used in the proof of the theorem is from [**JRB:71**]. An elementary proof of Corollary 6.6.27 can be found in [**WRB/RHB:71**].

Exercises

6.6.1 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a finite-dimensional normed $\mathbb{F}$-vector space. Show that a subspace $\mathsf{U} \subseteq \mathsf{V}$ is dense in V if and only if $\mathsf{U} = \mathsf{V}$. Point out which parts of your argument are not generally valid when V is infinite-dimensional.

6.6.2 Let $(\mathsf{V}, \|\cdot\|)$ be a normed vector space and let $A, B, C \subseteq \mathsf{V}$ be subsets with $A \subseteq B \subseteq C$. Show that if A is dense in B and if B is dense in C then A is dense in C.

6.6.3 Consider Example 6.6.22. On the subspace U (resp. V) denote the restriction of $\|\cdot\|_2$ by $\|\cdot\|_{\mathsf{U}}$ (resp. $\|\cdot\|_{\mathsf{V}}$). By Proposition 6.3.4 the normed vector space $\mathsf{U}_1 \oplus \mathsf{U}_2$ is complete. But in Example 6.6.22 we showed that $\mathsf{U}_1 \oplus \mathsf{U}_2$ is not a closed subspace of $\ell^2(\mathbb{F})$ and so is not complete by Corollary 6.6.17.

Why are these conclusions not in contradiction?

6.6.4 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space. If $\mathsf{U} \subseteq \mathsf{V}$ is a subspace, show that $\operatorname{cl}(\mathsf{U})$ is a subspace.

Section 6.7

Examples of Banach spaces

In this section we consider some of the common Banach spaces we will encounter in these volumes. As has already been mentioned, these examples serve as more than just an illustration of the concept of a Banach space; the examples are of great interest *per se*. Many of the examples are interconnected in that there is a very general example that contains simpler ones as a subcase. Logically, the proper way to present such examples is to give the most general construction first, and then provide the particular situations as following from the general. However, this method of presentation has serious defect that we are often most interested in the simpler situation, and a purely logical presentation would require the reader to understand some unnecessary abstraction. Therefore, we present our examples in order from the most particular to the most general. This has the drawback of being repetitive, but the advantage that a reader will not have to absorb a degree of abstraction that is not needed in the simpler examples.

Do I need to read this section? As we have said, some of the examples in this section are crucial in understanding a lot of the applied material that will follow. As the very least the reader should understand the spaces $\mathsf{L}^p(I;\mathbb{F})$ and $\ell^p(\mathbb{F})$. Some of the other examples can perhaps be omitted on a first reading, and covered when needed. •

6.7.1 The p-norms on $\mathbb{F}^n$

Let $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$. Let us begin our presentation with the simplest situation of a class of norms on a finite-dimensional $\mathbb{F}$-vector space. We are interested in a concrete collection of norms on the vector space $\mathbb{F}^n$. Specifically, for $p\in[1,\infty]$ we define a norm $\|\cdot\|_p$ on $\mathbb{F}^n$ by

$$\|(v_1,\dots,v_n)\|_p=\begin{cases}\left(\sum_{j=1}^n|v_j|^p\right)^{1/p}, & p\in[1,\infty),\\ \max\{|v_1|,\dots,|v_n|\}, & p=\infty.\end{cases}$$

That this is a norm for $p\in\{1,\infty\}$ has already been shown in Examples 6.1.3–3 and 6.1.3–4. In order to show that $\|\cdot\|_p$ is a norm for $p\in[1,\infty)$, the only nontrivial verification is of the triangle inequality. We verify this by using the following lemma.

6.7.1 Lemma (Hölder's inequality) *If* $\mathrm{a}_1,\dots,\mathrm{a}_\mathrm{n},\mathrm{b}_1,\dots,\mathrm{b}_\mathrm{n}\in\mathbb{R}_{\geq0}$ *and if* $\mathrm{p}\in(1,\infty)$ *then*

$$\sum_{\mathrm{j}=1}^{\mathrm{n}}\mathrm{a}_\mathrm{j}\mathrm{b}_\mathrm{j}\leq\Big(\sum_{\mathrm{j}=1}^{\mathrm{n}}\mathrm{a}_\mathrm{j}^\mathrm{p}\Big)^{1/\mathrm{p}}\Big(\sum_{\mathrm{j}=1}^{\mathrm{n}}\mathrm{b}_\mathrm{j}^{\mathrm{p}'}\Big)^{1/\mathrm{p}'},$$

where $\frac{1}{\mathrm{p}}+\frac{1}{\mathrm{p}'}=1$. *Moreover, equality holds if and only if* $(\mathrm{a}_1^\mathrm{p},\dots,\mathrm{a}_\mathrm{n}^\mathrm{p})$ *and* $(\mathrm{b}_1^{\mathrm{p}'},\dots,\mathrm{b}_\mathrm{n}^{\mathrm{p}'})$ *are collinear.*

Proof We first prove a lemma.

1 Sublemma *If* $\mathrm{a}, \mathrm{b} \in \mathbb{R}_{\geq 0}$ *and if* $\alpha \in (0,1)$ *then*

$$\mathrm{a}^{\alpha}\mathrm{b}^{1-\alpha} \leq \alpha\mathrm{a} + (1-\alpha)\mathrm{b},$$

and equality holds if and only if $\mathrm{a} = \mathrm{b}$.

Proof If $a = b$ then both sides of the inequality are equal to a, and so the result holds in this case. Thus we consider the case when $a \neq b$. Since the desired inequality is symmetric with respect to a and b we can assume that $b > a$ without loss of generality. Consider the function $f\colon [a,b] \to \mathbb{R}$ defined by $f(x) = x^{1-\alpha}$. By the Mean Value Theorem there exists $c \in (a,b)$ such that

$$f'(c) = (1-\alpha)c^{-\alpha} = \frac{f(b)-f(a)}{b-a} = \frac{b^{1-\alpha}-a^{1-\alpha}}{b-a}.$$

Thus $b^{1-\alpha} - a^{1-\alpha} = (b-a)(1-\alpha)c^{-\alpha}$. Since $\alpha \in (0,1)$ and since $c > a$ it follows that $c^{-\alpha} < a^{-\alpha}$. Therefore,

$$\begin{aligned} & b^{1-\alpha} - a^{1-\alpha} < (b-a)(1-\alpha)a^{-\alpha}, \\ \implies\quad & a^{\alpha}b^{1-\alpha} - a < (b-a)(1-\alpha) \\ \implies\quad & a^{\alpha}b^{1-\alpha} < \alpha a + (1-\alpha)b. \end{aligned}$$

Since this inequality is strict for $b > a$ the result follows. ▼

Let us denote $\alpha = \frac{1}{p}$ and $\beta = \frac{1}{p'} = 1 - \alpha$. Define $a'_j = a_j^{1/\alpha}$ and $b'_j = b_j^{1/\beta}$ and suppose initially that $\sum_{j=1}^n a'_j = 1$ and $\sum_{j=1}^n b'_j = 1$. By Sublemma 1 we have

$$\begin{aligned} & (a'_j)^{\alpha}(b'_j)^{\beta} \leq \alpha a'_j + \beta b'_j, \qquad j \in \{1,\dots,n\}, \\ \implies\quad & \sum_{j=1}^n ((a'_j)^{\alpha}(b'_j)^{\beta}) \leq \sum_{j=1}^n (\alpha a'_j + \beta b'_j) = \alpha + \beta = 1 = \Bigl(\sum_{j=1}^n a'_j\Bigr)^{\alpha}\Bigl(\sum_{j=1}^n b'_j\Bigr)^{\beta}, \\ \implies\quad & \sum_{j=1}^n a_j b_j \leq \Bigl(\sum_{j=1}^n a_j^p\Bigr)^{1/p}\Bigl(\sum_{j=1}^n b_j^{p'}\Bigr)^{1/p'}, \end{aligned}$$

with equality holding if and only if $a'_j = b'_j$, $j \in \{1,\dots,n\}$. This gives inequality in the sublemma when $\sum_{j=1}^n a'_j = 1$ and $\sum_{j=1}^n b'_j = 1$. If these relations do not hold then we have $\sum_{j=1}^n a'_j = \lambda$ and $\sum_{j=1}^n b'_j = \mu$ for some $\lambda, \mu \in \mathbb{R}_{\geq 0}$. Since the inequality is clearly equality if either $\lambda = 0$ or $\mu = 0$, we can suppose that $\lambda, \mu \in \mathbb{R}_{>0}$ without loss of generality. We can then write $a''_j = \frac{1}{\lambda}a'_j$ and $b''_j = \frac{1}{\mu}b'_j$ for $j \in \{1,\dots,n\}$ so that $\sum_{j=1}^n a''_j = \sum_{j=1}^n b''_j = 1$. Then

$$\begin{aligned} \sum_{j=1}^n a_j b_j &= \sum_{j=1}^n (a'_j)^{\alpha}(b'_j)^{\beta} = \lambda^{\alpha}\mu^{\beta}\sum_{j=1}^n (a''_j)^{\alpha}(b''_j)^{\beta} \\ &\leq \lambda^{\alpha}\mu^{\beta}\Bigl(\sum_{j=1}^n a''_j\Bigr)^{\alpha}\Bigl(\sum_{j=1}^n b''_j\Bigr)^{\beta} = \Bigl(\sum_{j=1}^n a'_j\Bigr)^{\alpha}\Bigl(\sum_{j=1}^n b'_j\Bigr)^{\beta} \\ &= \Bigl(\sum_{j=1}^n a_j^p\Bigr)^{1/p}\Bigl(\sum_{j=1}^n b_j^{p'}\Bigr)^{1/p'}, \end{aligned}$$

giving the desired inequality. Moreover, from our previous computations, equality holds if and only if $a''_j = b''_j$, $j \in \{1,\dots,n\}$. This, in turn, holds if and only if $\mu a'_j = \lambda b'_j$, $j \in \{1,\dots,n\}$. In turn, this holds if and only if

$$\mu a_j^p = \lambda b_j^{p'}, \qquad j \in \{1,\dots,n\},$$

which is the result. ■

6.7.2 Notation (Conjugate index) For $p \in (1,\infty)$ the number $p' \in (1,\infty)$ such that $\frac{1}{p} + \frac{1}{p'} = 1$ is called the ***conjugate index*** for the index p. As we shall see, principally in Section **??**, the conjugate index plays a surprisingly important rôle, although at this point it comes up as something of a conjurer's trick. Note that when $p = 2$ we have $p' = 2$. This in the important special case when the norm is derived from an inner product. We hope that the reader is tantalised at this moment. •

A variant of Hölder's inequality holds when $p = 1$, and we refer to Exercise 6.7.1 for this.

We next prove the useful Minkowski inequality.

6.7.3 Lemma (Minkowski's inequality) *If* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$, *if* $\mathrm{a}_1,\dots,\mathrm{a}_n,\mathrm{b}_1,\dots,\mathrm{b}_n \in \mathbb{F}$, *and if* $\mathrm{p} \in [1,\infty)$ *then*

$$\Big(\sum_{j=1}^{n}|\mathrm{a}_j + \mathrm{b}_j|^{\mathrm{p}}\Big)^{1/\mathrm{p}} \le \Big(\sum_{j=1}^{n}|\mathrm{a}_j|^{\mathrm{p}}\Big)^{1/\mathrm{p}} + \Big(\sum_{j=1}^{n}|\mathrm{b}_j|^{\mathrm{p}}\Big)^{1/\mathrm{p}}.$$

Moreover, equality holds if and only if the following conditions hold:

(i) $\mathrm{p} = 1$*: for each* $\mathrm{j} \in \{1,\dots,\mathrm{n}\}$ *there exists* $\alpha_\mathrm{j},\beta_\mathrm{j} \in \mathbb{R}_{\ge 0}$*, not both zero, such that* $\alpha_\mathrm{j}\mathrm{a}_\mathrm{j} = \beta_\mathrm{j}\mathrm{b}_\mathrm{j}$*;*

(ii) $\mathrm{p} \in (1,\infty)$*: there exists* $\alpha,\beta \in \mathbb{R}_{\ge 0}$*, not both zero, such that* $\alpha\mathrm{a}_\mathrm{j} = \beta\mathrm{b}_\mathrm{j}$ *for every* $\mathrm{j} \in \{1,\dots,\mathrm{n}\}$.

Proof The first part of the lemma has been proved for $p = 1$ in Example 6.1.3–3. Let us also prove the second part of the lemma for $p = 1$. First of all, it is easy to check that (i) is sufficient for equality in the Minkowski inequality. For the converse, note that, no matter whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, equality holds in the triangle inequality $|a+b| \le |a| + |b|$, $a,b \in \mathbb{F}$, if and only if there exists $\alpha,\beta \in \mathbb{R}_{\ge 0}$, not both zero, such that $\alpha a = \beta b$. The reader not seeing this is encouraged to do the elementary geometry needed to verify this. From this observation,

$$\sum_{j=1}^{n}|a_j + b_j| = \sum_{j=1}^{n}|a_j| + \sum_{j=1}^{n}|b_j|$$

if and only if (i) holds.

Since the case of $p = 1$ has already been proved, we consider $p \in (1, \infty)$. We compute, using Lemma 6.7.1,

$$\begin{aligned}\sum_{j=1}^n |a_j + b_j|^p &= \sum_{j=1}^n |a_j + b_j||a_j + b_j|^{p-1} \\ &\leq \sum_{j=1}^n |a_j||a_j + b_j|^{p-1} + \sum_{j=1}^n |b_j||a_j + b_j|^{p-1} \\ &\leq \Big(\sum_{j=1}^n |a_j|^p\Big)^{1/p}\Big(\sum_{j=1}^n |a_j + b_j|^p\Big)^{1/p'} + \Big(\sum_{j=1}^n |b_j|^p\Big)^{1/p}\Big(\sum_{j=1}^n |a_j + b_j|^p\Big)^{1/p'} \\ &= \Big(\Big(\sum_{j=1}^n |a_j|^p\Big)^{1/p} + \Big(\sum_{j=1}^n |b_j|^p\Big)^{1/p}\Big)\Big(\sum_{j=1}^n |a_j + b_j|^p\Big)^{1/p'}\end{aligned}$$

from which we deduce, using the fact that $\frac{1}{p} = 1 - \frac{1}{p'}$,

$$\Big(\sum_{j=1}^n |a_j + b_j|^p\Big)^{1/p} \leq \Big(\sum_{j=1}^n |a_j|^p\Big)^{1/p} + \Big(\sum_{j=1}^n |b_j|^p\Big)^{1/p},$$

as desired. By considering where the possible inequality is introduced in the preceding computation, and in view of Lemma 6.7.1, equality in the statement of the sublemma holds if and only if

1. for each $j \in \{1, \dots, n\}$ there exists $\alpha_j, \beta_j \in \mathbb{R}_{\geq 0}$, not both zero, such that $\alpha_j a_j = \beta_j b_j$ and
2. both $(|a_1|^p, \dots, |a_n|^p)$ and $(|b_1|^p, \dots, |b_n|^p)$ are collinear with $(|a_1 + b_1|^p, \dots, |a_n + b_n|^p)$.

The second of these conditions is equivalent to the existence of $\alpha, \lambda \in \mathbb{R}_{\geq 0}$, not both zero, and $\beta, \mu \in \mathbb{R}_{\geq 0}$, not both zero, such that

$$\alpha|a_j|^p = \lambda|a_j + b_j|^p, \quad \beta|b_j|^p = \mu|a_j + b_j|^p.$$

We consider a few cases.

1. $a_j, b_j \neq 0$ for every j: In this case we must have $\alpha_j, \beta_j \in \mathbb{R}_{>0}$. Then $a_j = \delta_j b_j$ for $\delta_j = \frac{\beta_j}{\alpha_j} \in \mathbb{R}_{>0}$. We can then solve for δ_j to give $\delta_j = \frac{\lambda}{\alpha - \lambda}$. Note that $\alpha \neq \lambda$ since $b_j \neq 0$. This gives
$$(\alpha - \lambda)a_j = \lambda b_j$$
for every $j \in \{1, \dots, n\}$, giving the result in this case.
2. $a_j, b_j \neq 0$ for some $j \in \{1, \dots, n\}$: In this case, whenever $a_j, b_j \neq 0$ the argument from the previous case gives
$$(\alpha - \lambda)a_j = \lambda b_j$$
Now we consider some subcases, taking into account that a_j and/or b_j might be zero for some j.
   (a) $a_j = 0$, $b_j \neq 0$: In this case we have $\lambda = \beta_j = 0$ and $\beta = \mu$. It, therefore, holds that
   $$(\alpha - \lambda)a_j = \lambda b_j.$$

(b) $a_j \neq 0, b_j = 0$: In this case $\mu = \alpha_j = 0$ and $\alpha = \lambda$. It, therefore, holds that

$$(\alpha - \lambda)a_j = \lambda b_j.$$

3. $a_j = 0$ for all $j \in \{1, \dots, n\}$: In this case we have, for any $\alpha \in \mathbb{R}_{>0}$ and with $\beta = 0$,

$$\alpha a_j = \beta b_j$$

for all $j \in \{1, \dots, n\}$.

4. $b_j = 0$ for all $j \in \{1, \dots, n\}$: In this case we have, for any $\beta \in \mathbb{R}_{>0}$ with $\alpha = 0$,

$$\alpha a_j = \beta b_j$$

for all $j \in \{1, \dots, n\}$.

The upshot of the preceding monotony is that condition (ii) holds when equality in the Minkowski inequality holds. ■

There is another version of the Minkowski inequality that is sometimes useful. We call this the "integral version" of the Minkowski inequality for reasons that are best made clear in ***missing stuff***.

6.7.4 Lemma (Integral version of Minkowski's inequality) ***missing stuff*** *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *if* $\mathrm{a_{jk}} \in \mathbb{F}$, $\mathrm{j} \in \{1, \dots, \mathrm{m}\}$, $\mathrm{k} \in \{1, \dots, \mathrm{n}\}$, *and if* $\mathrm{p} \in [1, \infty)$ *then*

$$\Big(\sum_{\mathrm{j}=1}^{\mathrm{m}} \Big|\sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{a_{jk}}\Big|^{\mathrm{p}}\Big)^{1/\mathrm{p}} \leq \sum_{\mathrm{k}=1}^{\mathrm{n}} \Big(\sum_{\mathrm{j}=1}^{\mathrm{m}} |\mathrm{a_{jk}}|^{\mathrm{p}}\Big)^{1/\mathrm{p}}.$$

Moreover, equality holds if and only if there exists $\mathrm{b_1}, \dots, \mathrm{b_m}, \mathrm{c_1}, \dots, \mathrm{c_n} \in \mathbb{F}$ *such that* $\mathrm{a_{jk}} = \mathrm{b_j c_k}$.

Proof For $p = 1$ we have

$$\sum_{j=1}^{m} \Big|\sum_{k=1}^{n} a_{jk}\Big| \leq \sum_{j=1}^{m} \Big(\sum_{k=1}^{n} |a_{jk}|\Big) = \sum_{k=1}^{n} \Big(\sum_{j=1}^{m} |a_{jk}|\Big),$$

giving the result in this case.

Now let $p \in (1, \infty)$. Here we compute

$$\begin{aligned}
\sum_{j=1}^{m} \Big|\sum_{k=1}^{n} a_{jk}\Big|^{p} &= \sum_{j=1}^{m} \Big(\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p-1}\Big)\Big(\Big|\sum_{l=1}^{n} a_{jl}\Big|\Big) \\
&\leq \sum_{j=1}^{m} \Big(\sum_{l=1}^{n} \Big(|a_{jl}| \Big|\sum_{k=1}^{n} a_{jk}\Big|^{p-1}\Big)\Big) \\
&= \sum_{l=1}^{n} \Big(\sum_{j=1}^{m} \Big(|a_{jl}| \Big|\sum_{k=1}^{n} a_{jk}\Big|^{p-1}\Big)\Big),
\end{aligned}$$

swapping the order of summation in the last step. Now let $p' = \frac{p}{p-1}$ be the conjugate index. Now, by Hölder's inequality,

$$\sum_{j=1}^{m}\Big(|a_{jl}|\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p-1}\Big) \le \Big(\sum_{j=1}^{m}|a_{jl}|^{p}\Big)^{1/p}\Big(\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p'(p-1)}\Big)^{1/p'}$$
$$= \Big(\sum_{j=1}^{m}|a_{jl}|^{p}\Big)^{1/p}\Big(\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p}\Big)^{1/p'}.$$

Substituting this last relation into the preceding equation yields

$$\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p} \le \sum_{l=1}^{n}\Big(\Big(\sum_{j=1}^{m}|a_{jl}|^{p}\Big)^{1/p}\Big(\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p}\Big)^{1/p'}\Big)$$
$$= \Big(\sum_{l=1}^{n}\Big(\sum_{j=1}^{m}|a_{jl}|^{p}\Big)^{1/p}\Big)\Big(\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p}\Big)^{1/p'}.$$

Now we note that the lemma is obviously true when

$$\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p} = 0.$$

So we suppose that this quantity is nonzero and divide the above-derived inequality

$$\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p} \le \Big(\sum_{l=1}^{n}\Big(\sum_{j=1}^{m}|a_{jl}|^{p}\Big)^{1/p}\Big)\Big(\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p}\Big)^{1/p'}$$

by

$$\Big(\sum_{j=1}^{m}\Big|\sum_{k=1}^{n} a_{jk}\Big|^{p}\Big)^{1/p'},$$

which gives the desired inequality after noting that p' is conjugate to p. ■

From Minkowski's inequality we immediately have the following result.

6.7.5 Proposition (($\mathbb{F}^n, \|\cdot\|_p$) is a Banach space) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $p \in [1, \infty]$ *then* $(\mathbb{F}^n, \|\cdot\|_p)$ *is an* $\mathbb{F}$*-Banach space.*

Moreover, we know from Theorem 6.1.15 that the norms $\|\cdot\|_p$ are equivalent. One can then wonder at why one would not just choose one of these norms and be done with it. There are at least two reasons why not.

1. Sometimes one norm is more convenient than another.
2. The finite-dimensional setting provides an opportunity to begin to understand the rôle p in how the norms "look." These sorts of p-norms will come up in increasingly more abstract settings, and the finite-dimensional example gives some useful intuition.

Along the lines of using the finite-dimensional setting to provide some intuition for more complicated ideas that will arise later, let us consider a variant of the p-norm for $p \in (0,1)$. The definition is the same. For $p \in (0,1)$ we define

$$\|(v_1,\dots,v_n)\|_p = \Big(\sum_{j=1}^{n} |v_j|^p\Big)^{1/p}.$$

The function $v \mapsto \|v\|_p$ clearly has the positivity and homogeneity properties needed for a norm. What we lose is the triangle inequality. Indeed, for $p \in (0,1)$ we have the following results which mirror Lemmata 6.7.1 and 6.7.3.

6.7.6 Lemma (Hölder's inequality for p ∈ (0,1)) *If* $\mathrm{a}_1,\dots,\mathrm{a}_n,\mathrm{b}_1,\dots,\mathrm{b}_n \in \mathbb{R}_{\geq 0}$ *and if* $\mathrm{p} \in (0,1)$ *then*

$$\sum_{j=1}^{n} \mathrm{a}_j\mathrm{b}_j \geq \Big(\sum_{j=1}^{n} \mathrm{a}_j^{\mathrm{p}}\Big)^{1/\mathrm{p}} \Big(\sum_{j=1}^{n} \mathrm{b}_j^{\mathrm{p}'}\Big)^{1/\mathrm{p}'},$$

where $\frac{1}{\mathrm{p}} + \frac{1}{\mathrm{p}'} = 1$.

Proof Let $q = p^{-1}$ so that $q \in (1,\infty)$ and define $c_j = b_j^{-1/q}$ and $d_j = a_j^{1/q} b_j^{1/q}$, $j \in \{1,\dots,n\}$. Let q' satisfy $\frac{1}{q} + \frac{1}{q'} = 1$. Then one shows directly that $c_j^{q'} = b_j^{p'}$ and $a_j^p = c_j d_j$, $j \in \{1,\dots,n\}$. Then we have, using Lemma 6.7.1,

$$\sum_{j=1}^{n} a_j^p = \sum_{j=1}^{n} c_j d_j \leq \Big(\sum_{j=1}^{n} d_j^q\Big)^{1/q} \Big(\sum_{j=1}^{n} c_j^{q'}\Big)^{1/q'} = \Big(\sum_{j=1}^{n} a_j b_j\Big)^{p} \Big(\sum_{j=1}^{n} b_j^{p'}\Big)^{1/q'},$$

from which we deduce that

$$\sum_{j=1}^{n} a_j b_j \geq \Big(\sum_{j=1}^{n} a_j^p\Big)^{1/p} \Big(\sum_{j=1}^{n} b_j^{p'}\Big)^{-1/(q'p)},$$

from which the result follows since $-\frac{1}{q'p} = \frac{1}{p'}$. ■

6.7.7 Lemma (Minkowski's inequality for p ∈ (0,1)) *If* $\mathrm{a}_1,\dots,\mathrm{a}_n,\mathrm{b}_1,\dots,\mathrm{b}_n \in \mathbb{R}_{\geq 0}$ *and if* $\mathrm{p} \in (0,1)$ *then*

$$\Big(\sum_{j=1}^{n} (\mathrm{a}_j + \mathrm{b}_j)^{\mathrm{p}}\Big)^{1/\mathrm{p}} \geq \Big(\sum_{j=1}^{n} \mathrm{a}_j^{\mathrm{p}}\Big)^{1/\mathrm{p}} + \Big(\sum_{j=1}^{n} \mathrm{b}_j^{\mathrm{p}}\Big)^{1/\mathrm{p}},$$

Proof This follows from Lemma 6.7.6 using the same sequence of computations used in proving that Lemma 6.7.3 follows from Lemma 6.7.1. ■

In Figure 6.4 we depict the boundaries of the balls $\mathsf{B}_p(1,\mathbf{0})$ in $\mathbb{R}^2$. The main point is that the balls are convex of and only if $p \in [1,\infty)$. In the present finite-dimensional setting this has no consequences. One can define a topology on $\mathbb{F}^n$ as being generated by the open balls, even though they are not convex. This topology is equivalent to the standard topology (one can see this by applying ***missing stuff***), and so all the usual notions of convergence, continuity, etc., carry over to this case. However, when we generalise this to infinite-dimensions, it turns out that the lack of convexity causes problems. For example, in ***missing stuff*** we shall see that the lack of convexity causes the topological dual to consist only of the zero functional.

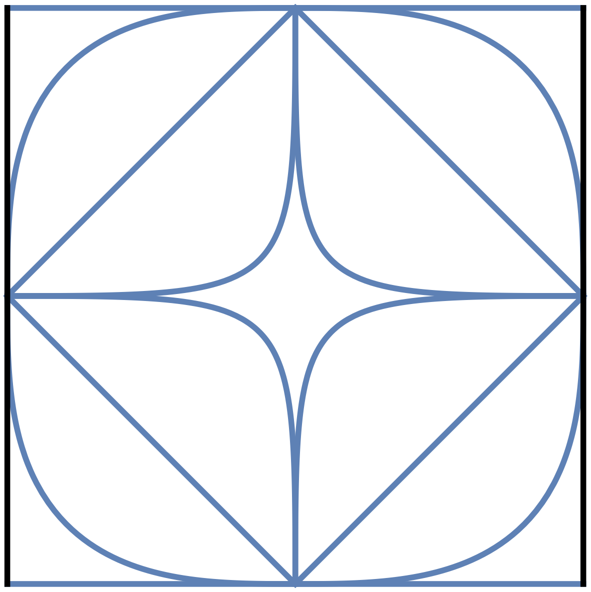


Figure 6.4 The unit spheres for the (if $p > 1$, at least) norms $\|\cdot\|_2$ on $\mathbb{R}^2$ (shown are, from inside to out, $p \in \{1/3, 1, 3, \infty\}$)

6.7.2 Banach spaces of sequences

Among the more important classes of Banach spaces we will encounter are those that are sequences characterised by certain summability properties. As we shall expound on in detail in ***missing stuff***, such Banach spaces are models for discrete time- and frequency-domain representations of signals. Here we are merely interested in some basic definitions and properties.

The most fundamental Banach space of sequences are those that are bounded.

6.7.8 Definition ($\ell^\infty(\mathbb{F})$) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Define a subspace $\ell^\infty(\mathbb{F})$ of $\mathbb{F}^{\mathbb{Z}_{>0}}$ by

$$\ell^\infty(\mathbb{F}) = \{(a_j)_{j\in\infty} \mid \text{ there exists } M \in \mathbb{R}_{>0} \text{ such that } |a_j| \leq M,\ j \in \mathbb{Z}_{>0}\}$$

and define

$$\|(a_j)_{j\in\mathbb{Z}_{>0}}\|_\infty = \sup\{|a_j| \mid j \in \mathbb{Z}_{>0}\}$$

for $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^\infty(\mathbb{F})$. •

Thus $\ell^\infty(\mathbb{F})$ consists of the set of bounded sequences in $\mathbb{F}$ and $\|\cdot\|_\infty$ is the least upper bound for the terms in the sequence.

6.7.9 Theorem ($\ell^\infty(\mathbb{F})$ is a Banach space) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ then $(\ell^\infty(\mathbb{F}), \|\cdot\|_\infty)$ is an $\mathbb{F}$-Banach space.*

Proof The only not entirely trivial norm property to verify for $\|\cdot\|_\infty$ is the triangle inequality:

$$\begin{aligned}\|(a_j)_{j\in\mathbb{Z}_{>0}} + (b_j)_{j\in\mathbb{Z}_{>0}}\|_\infty &= \sup\{|a_j + b_j| \mid j \in \mathbb{Z}_{>0}\}\\ &\leq \sup\{|a_j| + |b_j| \mid j \in \mathbb{Z}_{>0}\}\\ &= \sup\{|a_j| \mid j \in \mathbb{Z}_{>0}\} + \sup\{|b_j| \mid j \in \mathbb{Z}_{>0}\}\\ &= \|(a_j)_{j\in\mathbb{Z}_{>0}}\|_\infty + \|(b_j)_{j\in\mathbb{Z}_{>0}}\|_\infty,\end{aligned}$$

where we have used Proposition 2.2.27.

Now let us verify that $(\ell^\infty(\mathbb{F}), \|\cdot\|_\infty)$ is complete. We let $((a_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\ell^\infty(\mathbb{F})$. We claim that, for each $j \in \mathbb{Z}_{>0}$, $(a_j^{(l)})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{F}$. To see this, let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left\|(a_j^{(l)})_{j\in\mathbb{Z}_{>0}} - (a_j^{(k)})_{j\in\mathbb{Z}_{>0}}\right\|_\infty < \epsilon$$

for $k, l \geq N$. Then, by definition of $\|\cdot\|_\infty$,

$$\left|a_j^{(l)} - a_j^{(k)}\right| < \epsilon$$

for $k, l \geq N$ and for $j \in \mathbb{Z}_{>0}$. Thus $(a_j^{(l)})_{l\in\mathbb{Z}_{>0}}$ is indeed a Cauchy sequence, and so converges to some $a_j \in \mathbb{F}$. We now claim that the sequence $((a_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ converges to $(a_j)_{j\in\mathbb{Z}_{>0}}$. To see this, let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left\|(a_j^{(l)})_{j\in\mathbb{Z}_{>0}} - (a_j^{(k)})_{j\in\mathbb{Z}_{>0}}\right\|_\infty < \tfrac{\epsilon}{2}$$

for $k, l \geq N$. Thus

$$\left|a_j^{(l)} - a_j^{(l)}\right| < \tfrac{\epsilon}{2}, \qquad k, l \geq N.$$

Now, for fixed $j \in \mathbb{Z}_{>0}$, let $N' \in \mathbb{Z}_{>0}$ be sufficiently large that $\left|a_j^{(k)} - a_j\right| < \frac{\epsilon}{2}$ for $k \geq N'$. In this case, if $l \geq N$ and $k \geq \max\{N, N'\}$, we have

$$\left|a_j^{(l)} - a_j\right| \leq \left|a_j^{(l)} - a_j^{(k)}\right| + \left|a_j^{(k)} - a_j\right| < \epsilon.$$

Since this holds for each $j \in \mathbb{Z}_{>0}$ we have

$$\left\|(a_j^{(l)})_{j\in\mathbb{Z}_{>0}} - (a_j)_{j\in\mathbb{Z}_{>0}}\right\|_\infty \leq \epsilon,$$

as desired. ■

One property of $\ell^\infty(\mathbb{F})$ that makes it different than some of the other Banach spaces we consider is the following.

6.7.10 Proposition ($\ell^\infty(\mathbb{F})$ is not separable) *For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, the Banach space $(\ell^\infty(\mathbb{F}), \|\cdot\|_\infty)$ is not separable.*

Proof Let $\mathscr{U}$ be the collection of sequences $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^\infty(\mathbb{F})$ such that $a_j \in \{-1, 1\}$, $j \in \mathbb{Z}_{>0}$. It follows from Exercises **??**, **??**, and 2.1.4 that $\mathscr{U}$ is countable. Note that if $(a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{U}$ are distinct then

$$\|(a_j)_{j\in\mathbb{Z}_{>0}}\|_\infty = 1, \quad \|(a_j)_{j\in\mathbb{Z}_{>0}} - (b_j)_{j\in\mathbb{Z}_{>0}}\|_\infty = 2.$$

Let $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{U}$ and let $(b_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{B}(1,(a_j)_{j\in\mathbb{Z}_{>0}})$. By Exercise 6.1.3 we have

$$\begin{aligned} &\Big|\|(b_j)_{j\in\mathbb{Z}_{>0}}\| - \|(a_j)_{j\in\mathbb{Z}_{>0}}\|_\infty\Big| \le \Big\|(b_j)_{j\in\mathbb{Z}_{>0}} - (a_j)_{j\in\mathbb{Z}_{>0}}\Big\|_\infty \\ \implies\quad &\Big|\|(b_j)_{j\in\mathbb{Z}_{>0}}\| - 1\Big| \le 1 \\ \implies\quad &\Big\|(b_j)_{j\in\mathbb{Z}_{>0}}\Big\|_\infty \le 2. \end{aligned}$$

Thus $\mathsf{B}(1,(a_j)_{j\in\mathbb{Z}_{>0}}) \subseteq \mathsf{B}(2,0_{\ell^\infty(\mathbb{F})})$ for each $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{U}$. If

$$(a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{U}$$

and

$$(c_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{B}(1,(a_j)_{j\in\mathbb{Z}_{>0}}), \quad (d_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{B}(1,(b_j)_{j\in\mathbb{Z}_{>0}})$$

then

$$\begin{aligned} &\|(c_j)_{j\in\mathbb{Z}_{>0}} - (b_j)_{j\in\mathbb{Z}_{>0}}\|_\infty \\ &\qquad\qquad \ge \big|\|(c_j)_{j\in\mathbb{Z}_{>0}} - (a_j)_{j\in\mathbb{Z}_{>0}}\|_\infty - \|(a_j)_{j\in\mathbb{Z}_{>0}} - (b_j)_{j\in\mathbb{Z}_{>0}}\|_\infty\big| \ge 2 \end{aligned}$$

using Proposition **??**. Thus $(c_j)_{j\in\mathbb{Z}_{>0}} \notin \mathsf{B}(1,(b_j)_{j\in\mathbb{Z}_{>0}})$. One similarly shows that $(d_j)_{j\in\mathbb{Z}_{>0}} \notin \mathsf{B}(1,(a_j)_{j\in\mathbb{Z}_{>0}})$. This shows that $\mathsf{B}(2,0_{\ell^\infty(\mathbb{F})})$ contains the collection

$$\{\mathsf{B}(1,(a_j)_{j\in\mathbb{Z}_{>0}}) \mid (a_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{U}\}$$

of disjoint open balls. In particular, if $((b_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ is any countable subset of $\ell^\infty(\mathbb{F})$ then there is a countable or finite subset $((a_j^{(\alpha)})_{j\in\mathbb{Z}_{>0}})_{\alpha\in A}$ of $\mathscr{U}$ in which are contained all of the sequences $((b_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$. Note that

$$\mathrm{cl}\Big(((b_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}\Big) \subseteq \cup_{\alpha\in A}\overline{\mathsf{B}}(1,(a_j^{(\alpha)})_{j\in\mathbb{Z}_{>0}}).$$

Therefore, any of the set of balls

$$\{\mathsf{B}(1,(a_j)_{j\in\mathbb{Z}_{>0}}) \mid (a_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{U},\ (a_j)_{j\in\mathbb{Z}_{>0}} \ne (a_j^{(\alpha)})_{j\in\mathbb{Z}_{>0}},\ \alpha\in A\}$$

cannot lie in $\mathrm{cl}\Big(((b_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}\Big)$ which prohibits $((b_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ from being dense. ■

Now we begin looking at subspaces of $\ell^\infty(\mathbb{F})$. We begin with subspaces of sequences that converge.

6.7.11 Definition ($\mathsf{c}(\mathbb{F})$ and $\mathsf{c}_0(\mathbb{F})$) Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$. Define subspaces $\mathsf{c}(\mathbb{F})$ and $\mathsf{c}_0(\mathbb{F})$ of $\mathbb{F}^{\mathbb{Z}_{>0}}$ by

$$\mathsf{c}(\mathbb{F}) = \Big\{(a_j)_{j\in\mathbb{Z}_{>0}} \;\Big|\; \text{there exists } a\in\mathbb{F} \text{ such that } \lim_{j\to\infty} a_j = a\Big\}$$

and

$$\mathsf{c}_0(\mathbb{F}) = \Big\{(a_j)_{j\in\mathbb{Z}_{>0}} \;\Big|\; \lim_{j\to\infty} a_j = 0\Big\},$$

respectively. •

Note that by Propositions 2.3.23 and ***missing stuff*** it follows that $\mathsf{c}(\mathbb{F})$ and $\mathsf{c}_0(\mathbb{F})$ are subspaces. Moreover, by Propositions 2.3.3 and ***missing stuff*** it follows that $\mathsf{c}(\mathbb{F})$ and $\mathsf{c}_0(\mathbb{F})$ are subspaces of $\ell^\infty(\mathbb{F})$. The appropriate norm to use on the spaces of sequences $\mathsf{c}(\mathbb{F})$ and $\mathsf{c}_0(\mathbb{F})$ is the restriction of norm $\|\cdot\|_\infty$ on $\ell^\infty(\mathbb{F})$. We denote this norm simply by $\|\cdot\|_\infty$. With this norm our spaces of convergent sequences are Banach spaces.

6.7.12 Theorem (($\mathsf{c}(\mathbb{F}), \|\cdot\|_\infty$) and ($\mathsf{c}_0(\mathbb{F}), \|\cdot\|_\infty$) are Banach spaces) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *then* $(\mathsf{c}(\mathbb{F}), \|\cdot\|_\infty)$ *and* $(\mathsf{c}_0(\mathbb{F}), \|\cdot\|_\infty)$ *are* $\mathbb{F}$*-Banach spaces.*

Proof Let $((a_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathsf{c}(\mathbb{F})$. By Theorem 6.7.9 this means that the sequence converges to $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^\infty(\mathbb{F})$. Since each sequence $(a_j^{(l)})_{j\in\mathbb{Z}_{>0}}$ is in $\mathsf{c}(\mathbb{F})$ there exists $a^{(l)} \in \mathbb{F}$ such that $\lim_{j\to\infty} a_j^{(l)} = a^{(l)}$. We claim that $(a^{(l)})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. For $\epsilon \in \mathbb{R}_{>0}$ let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left\|(a_j^{(k)})_{j\in\mathbb{Z}_{>0}} - (a_j^{(l)})_{j\in\mathbb{Z}_{>0}}\right\|_\infty < \tfrac{\epsilon}{3},$$

which implies that

$$\left|a_j^{(k)} - a_j^{(l)}\right| < \tfrac{\epsilon}{3}, \qquad k, l \geq N,\ j \in \mathbb{Z}_{>0}.$$

Now let $k, l \geq N$ and let $j \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left|a_j^{(k)} - a^{(k)}\right| < \tfrac{\epsilon}{3}, \quad \left|a_j^{(l)} - a^{(k)}\right| < \tfrac{\epsilon}{3}.$$

Then

$$\left|a^{(k)} - a^{(l)}\right| \leq \left|a^{(k)} - a_j^{(k)}\right| + \left|a_j^{(k)} - a_j^{(l)}\right| + \left|a_j^{(l)} - a^{(l)}\right| < \epsilon.$$

As this holds for every $k, l \geq N$ it follows that $(a^{(l)})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{F}$. We denote its limit by a.

Finally we show that $\lim_{j\to\infty} a_j = a$, which shows that $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{c}(\mathbb{F})$. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N' \in \mathbb{Z}_{>0}$ be sufficiently large that $\left|a^{(k)} - a\right| < \frac{\epsilon}{3}$ for $k \geq N'$. Now fix $k \geq N'$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left|a_j - a_j^{(k)}\right| < \tfrac{\epsilon}{3}, \quad \left|a_j^{(k)} - a^{(k)}\right| < \tfrac{\epsilon}{3}, \qquad j \geq N.$$

Then, for $j \geq N$,

$$|a_j - a| \leq \left|a_j - a_j^{(k)}\right| + \left|a_j^{(k)} - a^{(k)}\right| + \left|a^{(k)} - a\right| < \epsilon,$$

which completes the proof that $(\mathsf{c}(\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space.

If $((a_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathsf{c}_0(\mathbb{F}) \subseteq \mathsf{c}(\mathbb{F})$ the above argument is easily modified to show that the limit sequence, denoted $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{c}(\mathbb{F})$ above is actually in $\mathsf{c}_0(\mathbb{F})$. The key point is that $a^{(l)} = 0$ for each $l \in \mathbb{Z}_{>0}$ and so $a = 0$ as well. Thus $(\mathsf{c}_0(\mathbb{F}), \|\cdot\|_\infty)$ is also a Banach space. ■

The Banach spaces $\mathsf{c}(\mathbb{F})$ an $\mathsf{c}_0(\mathbb{F})$ have the friendly property of being separable.

6.7.13 Proposition ($\mathsf{c}(\mathbb{F})$ and $\mathsf{c}_0(\mathbb{F})$ are separable) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *then the Banach spaces* $(\mathsf{c}(\mathbb{F}), \|\cdot\|_\infty)$ *and* $(\mathsf{c}_0(\mathbb{F}), \|\cdot\|_\infty)$ *are separable.*

Proof It suffices to prove the proposition for $\mathsf{c}(\mathbb{F})$. We first take the case when $\mathbb{F} = \mathbb{R}$. In this case, for $q \in \mathbb{Q}$, we let $\mathscr{D}_q(\mathbb{R})$ be the subset of $\mathsf{c}(\mathbb{R})$ consisting of sequences $(q_j)_{j\in\mathbb{Z}_{>0}}$ with $q_j \in \mathbb{Q}$, $j \in \mathbb{Z}_{>0}$, and such that $q_j = q$ for all j sufficiently large. We then take

$$\mathscr{D}(\mathbb{R}) = \cup_{q\in\mathbb{Q}} \mathscr{D}_q(\mathbb{R}).$$

We claim that $\mathscr{D}(\mathbb{R})$ is countable. We note that $\mathscr{D}_q(\mathbb{R})$ is a countable (indexed by $\mathbb{Z}_{>0}$) disjoint union of copies of $\mathbb{Q}$ and so is countable by Proposition **??**. Thus $\mathscr{D}(\mathbb{R})$ is a countable union of countable sets, and so is again countable by Proposition **??**. We should also show that $\mathscr{D}(\mathbb{R})$ is dense in $\mathsf{c}(\mathbb{R})$. Let $(a_j)_{j\in\mathbb{Z}_{>0}}$ and let $\epsilon \in \mathbb{R}_{>0}$. Suppose that $q \in \mathbb{Q}$ is such that

$$\Big|\lim_{j\to\infty} a_j - q\Big| < \epsilon$$

and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that $|a_j - q| < \epsilon$ for $j \geq N$. Now choose $q_1, \dots, q_N \in \mathbb{Q}$ such that $|a_j - q_j| < \epsilon$ for $j \in \{1, \dots, N\}$. Now define $(q_j)_{j\in\mathbb{Z}_{>0}}$ by asking that $q_j = q$ for $j > N$. Then $(q_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{D}(\mathbb{R})$ and

$$\|(a_j)_{j\in\mathbb{Z}_{>0}} - (q_j)_{j\in\mathbb{Z}_{>0}}\|_\infty < \epsilon.$$

Thus $\mathscr{D}(\mathbb{R})$ is dense in $\mathsf{c}(\mathbb{R})$.

For $\mathbb{F} = \mathbb{C}$ the procedure above can be duplicated by letting $\mathscr{D}(\mathbb{C})$ be the set of sequences $(q_j + \mathrm{i}r_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{c}(\mathbb{C})$ with $(q_j)_{j\in\mathbb{Z}_{>0}}, (r_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{D}(\mathbb{R})$. ■

The result has the following interesting corollary.

6.7.14 Corollary ($\mathsf{c}_0(\mathbb{F})$ is the completion of $\mathbb{F}_0^\infty$) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *then* $(\mathsf{c}_0(\mathbb{F}), \|\cdot\|_\infty)$ *is the completion of* $(\mathbb{F}_0^\infty, \|\cdot\|_\infty)$.

Proof Borrowing the notation from the proof of Proposition 6.7.13 we have

$$\mathscr{D}_0(\mathbb{F}) \subseteq \mathbb{F}_0^\infty \subseteq \mathsf{c}_0(\mathbb{F})$$

from which we deduce that

$$\mathsf{c}_0(\mathbb{F}) = \mathrm{cl}(\mathscr{D}_0(\mathbb{F})) \subseteq \mathrm{cl}(\mathbb{F}_0^\infty) \subseteq \mathrm{cl}(\mathsf{c}_0(\mathbb{F})) = \mathsf{c}_0(\mathbb{F}).$$

Therefore, $\mathrm{cl}(\mathbb{F}_0^\infty) = \mathsf{c}_0(\mathbb{F})$, as desired. ■

Now we consider Banach spaces of sequences which naturally use a different norm that the ∞-norm.

6.7.15 Definition ($\ell^p(\mathbb{F})$) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $p \in [1, \infty)$. Define a subspace $\ell^p(\mathbb{F})$ of $\mathbb{F}^{\mathbb{Z}_{>0}}$ by

$$\ell^p(\mathbb{F}) = \Big\{(a_j)_{j\in\mathbb{Z}_{>0}} \;\Big|\; \sum_{j=1}^\infty |a_j|^p < \infty\Big\}$$

and define

$$\|(a_j)_{j\in\mathbb{Z}_{>0}}\|_p = \Big(\sum_{j=1}^\infty |a_j|^p\Big)^{1/p}$$

for $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$. •

At this point it is not necessarily clear that $\ell^p(\mathbb{F})$ is actually a subspace of $\mathbb{F}^{\mathbb{Z}_{>0}}$, but we shall show shortly that it is, and is in fact a Banach space when equipped with $\|\cdot\|_p$ as a norm.

Let us give some properties of the function $\|\cdot\|_p$ analogous to Lemmata 6.7.1 and 6.7.3.

6.7.16 Lemma (Hölder's inequality) *If* $\mathrm{p} \in (1,\infty)$ *and if* $(\mathrm{a}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \ell^{\mathrm{p}}(\mathbb{F})$ *and* $(\mathrm{b}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \ell^{\mathrm{p}'}(\mathbb{F})$*, then*

$$\sum_{\mathrm{j}=1}^{\infty}|\mathrm{a}_\mathrm{j}\mathrm{b}_\mathrm{j}| \le \Big(\sum_{\mathrm{j}=1}^{\infty}|\mathrm{a}_\mathrm{j}|^{\mathrm{p}}\Big)^{1/\mathrm{p}}\Big(\sum_{\mathrm{j}=1}^{\infty}|\mathrm{b}_\mathrm{j}|^{\mathrm{p}'}\Big)^{1/\mathrm{p}'},$$

where $\frac{1}{\mathrm{p}} + \frac{1}{\mathrm{p}'} = 1$*. Moreover, equality holds if and only if* $(|\mathrm{a}_\mathrm{j}|^{\mathrm{p}})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *and* $(|\mathrm{b}_\mathrm{j}|^{\mathrm{p}'})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *are collinear.*

Proof For $N \in \mathbb{Z}_{>0}$, by Lemma 6.7.1 we have

$$\sum_{j=1}^{N}|a_jb_j| \le \Big(\sum_{j=1}^{N}|a_j|^p\Big)^{1/p}\Big(\sum_{j=1}^{N}|b_j|^{p'}\Big)^{1/p'} \le \Big(\sum_{j=1}^{\infty}|a_j|^p\Big)^{1/p}\Big(\sum_{j=1}^{\infty}|b_j|^{p'}\Big)^{1/p'}.$$

Thus

$$\sum_{j=1}^{\infty}|a_jb_j| = \lim_{N\to\infty}\sum_{j=1}^{N}|a_jb_j| \le \Big(\sum_{j=1}^{\infty}|a_j|^p\Big)^{1/p}\Big(\sum_{j=1}^{\infty}|b_j|^{p'}\Big)^{1/p'},$$

as desired.

For the final assertion of the lemma, first note that a direction computation shows that equality holds in the Hölder equality if $(|a_j|^p)_{j\in\mathbb{Z}_{>0}}$ and $(|b_j|^{p'})_{j\in\mathbb{Z}_{>0}}$ are collinear. For the converse, suppose that $(|a_j|^p)_{j\in\mathbb{Z}_{>0}}$ and $(|b_j|^{p'})_{j\in\mathbb{Z}_{>0}}$ are not collinear. Then there exists $N \in \mathbb{Z}_{>0}$ such that $(|a_1|^p,\dots,|a_N|^p)$ and $(|b_1|^{p'},\dots,|b_N|^{p'})$ are not collinear. By Lemma 6.7.1 we then have

$$\sum_{j=1}^{N}|a_jb_j| < \Big(\sum_{j=1}^{N}|a_j|^p\Big)^{1/p}\Big(\sum_{j=1}^{N}|bj|^{p'}\Big)^{1/p'}.$$

Since

$$\sum_{j=N+1}^{\infty}|a_jb_j| < \Big(\sum_{j=N+1}^{\infty}|a_j|^p\Big)^{1/p}\Big(\sum_{j=N+1}^{\infty}|bj|^{p'}\Big)^{1/p'}$$

it follows that equality cannot hold in the Hölder inequality. ■

A version of Hölder's inequality holds for $p = 1$ and we refer to Exercise 6.7.2 for this.

The Minkowski inequality also holds in this case.

6.7.17 Lemma (Minkowski's inequality) *If* $\mathrm{p} \in [1,\infty)$ *and if* $(\mathrm{a}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}}, (\mathrm{b}_\mathrm{j})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \ell^{\mathrm{p}}(\mathbb{F})$ *then*

$$\Big(\sum_{\mathrm{j}=1}^{\infty}|\mathrm{a}_\mathrm{j} + \mathrm{b}_\mathrm{j}|^{\mathrm{p}}\Big)^{1/\mathrm{p}} \le \Big(\sum_{\mathrm{j}=1}^{\infty}|\mathrm{a}_\mathrm{j}|^{\mathrm{p}}\Big)^{1/\mathrm{p}} + \Big(\sum_{\mathrm{j}=1}^{\infty}|\mathrm{b}_\mathrm{j}|^{\mathrm{p}}\Big)^{1/\mathrm{p}}.$$

Moreover, equality holds if and only if the following conditions hold:

(i) $\mathrm{p} = 1$*: for each* $\mathrm{j} \in \mathbb{Z}_{>0}$ *there exists* $\alpha_\mathrm{j}, \beta_\mathrm{j} \in \mathbb{R}_{\ge 0}$*, not both zero, such that* $\alpha_\mathrm{j}\mathrm{a}_\mathrm{j} = \beta_\mathrm{j}\mathrm{b}_\mathrm{j}$*;*

(ii) $\mathrm{p} \in (1,\infty)$*: there exists* $\alpha, \beta \in \mathbb{R}_{\ge 0}$*, not both zero, such that* $\alpha\mathrm{a}_\mathrm{j} = \beta\mathrm{b}_\mathrm{j}$ *for every* $\mathrm{j} \in \mathbb{Z}_{>0}$*.*

Proof Let $p \in [1,\infty)$ and let $(a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$. For each $N \in \mathbb{Z}_{>0}$

$$\Big(\sum_{j=1}^{N}|a_j+b_j|^p\Big)^{1/p} \le \Big(\sum_{j=1}^{N}|a_j|^p\Big)^{1/p} + \Big(\sum_{j=1}^{N}|b_j|^p\Big)^{1/p} \le \|(a_j)_{j\in\mathbb{Z}_{>0}}\|_p + \|(b_j)_{j\in\mathbb{Z}_{>0}}\|_p$$

by Lemma 6.7.3. Therefore,

$$\|(a_j)_{j\in\mathbb{Z}_{>0}} + (b_j)_{j\in\mathbb{Z}_{>0}}\|_p = \lim_{N\to\infty}\Big(\sum_{j=1}^{N}|a_j+b_j|^p\Big)^{1/p} \le \|(a_j)_{j\in\mathbb{Z}_{>0}}\|_p + \|(b_j)_{j\in\mathbb{Z}_{>0}}\|_p.$$

This shows that $(a_j)_{j\in\mathbb{Z}_{>0}} + (b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$.

An argument similar to that used in the last part of Lemma 6.7.16 can be used to prove the last assertion of the lemma. ■

The integral version of Minkowski's inequality also holds in this case.

6.7.18 Lemma (Integral version of Minkowski's inequality) *missing stuff If* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$, *if* $\mathrm{p} \in [1,\infty)$, *if* $\mathrm{a_{jk}} \in \mathbb{F}$, $\mathrm{j,k} \in \mathbb{Z}_{>0}$, *are such that* $(\mathrm{a_{jk}})_{\mathrm{j}\in\mathbb{Z}_{>0}} \in \ell^{\mathrm{p}}(\mathbb{F})$ *for every* $\mathrm{k} \in \mathbb{Z}_{>0}$ *and* $(\mathrm{a_{jk}})_{\mathrm{k}\in\mathbb{Z}_{>0}} \in \ell^{\mathrm{p}}(\mathbb{F})$ *for every* $\mathrm{j} \in \mathbb{Z}_{>0}$, *then*

$$\Big(\sum_{\mathrm{j}=1}^{\infty}\Big|\sum_{\mathrm{k}=1}^{\infty}\mathrm{a_{jk}}\Big|^{\mathrm{p}}\Big)^{1/\mathrm{p}} \le \sum_{\mathrm{k}=1}^{\infty}\Big(\sum_{\mathrm{j}=1}^{\infty}|\mathrm{a_{jk}}|^{\mathrm{p}}\Big)^{1/\mathrm{p}}.$$

Moreover, equality holds if and only if there exists $\mathrm{b_j,,c_k} \in \mathbb{F}$, $\mathrm{j,k} \in$]integerp, *such that* $\mathrm{a_{jk}} = \mathrm{b_jc_k}$.

Proof For $p = 1$ we have

$$\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big| \le \sum_{j=1}^{\infty}\Big(\sum_{k=1}^{\infty}|a_{jk}|\Big) = \sum_{k=1}^{\infty}\Big(\sum_{j=1}^{\infty}|a_{jk}|\Big),$$

giving the result in this case.

Now let $p \in (1,\infty)$. Here we compute

$$\begin{aligned}\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p &= \sum_{j=1}^{\infty}\Big(\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^{p-1}\Big)\Big(\Big|\sum_{l=1}^{\infty}a_{jl}\Big|\Big)\\ &\le \sum_{j=1}^{\infty}\Big(\sum_{l=1}^{\infty}\Big(|a_{jl}|\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^{p-1}\Big)\Big)\\ &= \sum_{l=1}^{\infty}\Big(\sum_{j=1}^{\infty}\Big(|a_{jl}|\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^{p-1}\Big)\Big),\end{aligned}$$

swapping the order of summation in the last step. Now let $p' = \frac{p}{p-1}$ be the conjugate index. Now, by Hölder's inequality,

$$\begin{aligned}\sum_{j=1}^{\infty}\Big(|a_{jl}|\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^{p-1}\Big) &\le \Big(\sum_{j=1}^{\infty}|a_{jl}|^{\infty}\Big)^{1/p}\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^{p'(p-1)}\Big)^{1/p'}\\ &= \Big(\sum_{j=1}^{\infty}|a_{jl}|^{p}\Big)^{1/p}\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^{p}\Big)^{1/p'}.\end{aligned}$$

Substituting this last relation into the preceding equation yields

$$\begin{aligned}\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p &\leq \sum_{l=1}^{\infty}\Big(\Big(\sum_{j=1}^{\infty}|a_{jl}|^p\Big)^{1/p}\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p\Big)^{1/p'}\Big)\\ &= \Big(\sum_{l=1}^{\infty}\Big(\sum_{j=1}^{\infty}|a_{jl}|^p\Big)^{1/p}\Big)\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p\Big)^{1/p'}.\end{aligned}$$

Now we note that the lemma is obviously true when

$$\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p = 0.$$

So we suppose that this quantity is nonzero and divide the above-derived inequality

$$\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p \leq \Big(\sum_{l=1}^{\infty}\Big(\sum_{j=1}^{\infty}|a_{jl}|^p\Big)^{1/p}\Big)\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p\Big)^{1/p'}$$

by

$$\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}a_{jk}\Big|^p\Big)^{1/p'},$$

which gives the desired inequality after noting that p' is conjugate to p. ■

Now we can prove that $\ell^p(\mathbb{F})$ is a Banach space.

6.7.19 Theorem ($(\ell^p(\mathbb{F}), \|\cdot\|_p)$ is a Banach space) *If* $\mathbb{F} \in \mathbb{R}, \mathbb{C}\}$ *and if* $p \in [1, \infty)$ *then* $(\ell^p(\mathbb{F}), \|\cdot\|_p)$ *is an* $\mathbb{F}$*-Banach space.*

Proof Let us first verify that $\ell^p(\mathbb{F})$ is a subspace. We first consider the case of $p = 1$. Let $(a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^1(\mathbb{F})$. By Lemma 6.7.17 we have $(a_j + b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^1(\mathbb{F})$. If $\alpha \in \mathbb{F}$ we have

$$\sum_{j=1}^{\infty}|\alpha a_j| = |\alpha|\sum_{j=1}^{\infty}|a_j|$$

by Proposition 2.4.30. Thus $\alpha(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^1(\mathbb{F})$, which shows that $\ell^1(\mathbb{F})$ is a subspace of $\mathbb{F}^{\mathbb{Z}_{>0}}$.

By Lemma 6.7.17, if $(a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$ then $(a_j + b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$. It is easy to see, just as for the case of $p = 1$, that $\alpha(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$ if $\alpha \in \mathbb{F}$ and if $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$, $p \in (1, \infty)$.

As we have shown the triangle inequality for $\|\cdot\|_p$ already in Lemma 6.7.17, and since the other norm properties for $\|\cdot\|_p$ hold trivially, it follows that $\ell^p(\mathbb{F})$ is a normed vector space. It remains to show that it is complete. Let $((a_j^{(l)})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\ell^p(\mathbb{F})$. We claim that the sequence $(a_j^{(l)})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence for each $j \in \mathbb{Z}_{>0}$. For every $\epsilon \in \mathbb{R}_{>0}$ there exists $N \in \mathbb{Z}_{>0}$ such that

$$\big\|(a_j^{(k)})_{j\in\mathbb{Z}_{>0}} - (a_j^{(l)})_{j\in\mathbb{Z}_{>0}}\big\|_p = \Big(\sum_{j=1}^{\infty}\big|a_j^{(k)} - a_j^{(l)}\big|^p\Big)^{1/p} < \epsilon.$$

Now let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that

$$\left\|(a_j^{(k)})_{j\in\mathbb{Z}_{>0}} - (b_j^{(l)})_{j\in\mathbb{Z}_{>0}}\right\|_p < \epsilon.$$

Then

$$\left|a_j^{(k)} - a_j^{(l)}\right|^p \leq \sum_{j=1}^{\infty}\left|a_j^{(k)} - a_j^{(l)}\right|^p < \epsilon^p,$$

giving $(a_j^{(l)})_{l\in\mathbb{Z}_{>0}}$ as a Cauchy sequence. Denote its limit by $a_j \in \mathbb{F}$. We next claim that $(a_j^{(l)})_{l\in\mathbb{Z}_{>0}}$ converges to $(a_j)_{j\in\mathbb{Z}_{>0}}$ in $\ell^p(\mathbb{F})$. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that

$$\left\|(a_j^{(l)})_{j\in\mathbb{Z}_{>0}} - (a_j^{(k)})_{j\in\mathbb{Z}_{>0}}\right\|_p < \tfrac{\epsilon}{2}, \qquad l,k \geq N.$$

For $n \in \mathbb{Z}_{>0}$ the sequence $((a_j^{(l)})_{j=1}^n)_{l\in\mathbb{Z}}$ converges to $(a_j)_{j=1}^n$ in $\mathbb{F}^n$ with respect to the norm $\|\cdot\|_p$ by Theorem 6.3.3. Thus there exists $N' \in \mathbb{Z}_{>0}$ such that

$$\Big(\sum_{j=1}^{n}\left|a_j^{(k)} - a_j\right|^p\Big)^{1/p} < \frac{\epsilon}{2}, \qquad k \geq N'.$$

Then, for $k \geq \max\{N, N'\}$,

$$\begin{aligned}\Big(\sum_{j=1}^{n}\left|a_j^{(l)} - a_j\right|^p\Big)^{1/p} &\leq \Big(\sum_{j=1}^{n}\left|a_j^{(l)} - a_j^{(k)}\right|^p\Big)^{1/p} + \Big(\sum_{j=1}^{n}\left|a_j^{(k)} - a_j\right|^p\Big)^{1/p} \\ &\leq \left\|(a_j^{(l)})_{j\in\mathbb{Z}_{>0}} - (a_j^{(k)})_{j\in\mathbb{Z}_{>0}}\right\|_p + \Big(\sum_{j=1}^{n}\left|a_j^{(k)} - a_j\right|^p\Big)^{1/p} < \epsilon.\end{aligned}$$

Now we have

$$\left\|(a_j^{(l)})_{j\in\mathbb{Z}_{>0}} - (a_j)_{j\in\mathbb{Z}_{>0}}\right\|_p = \lim_{n\to\infty}\Big(\sum_{j=1}^{n}\left|a_j^{(l)} - a_j\right|^p\Big)^{1/p} \leq \epsilon.$$

This gives convergence of $(a_j^{(l)})_{l\in\mathbb{Z}_{>0}}$ to $(a_j)_{j\in\mathbb{Z}_{>0}}$ in $\ell^p(\mathbb{F})$, as desired. ■

Let us show that, unlike $\ell^\infty(\mathbb{F})$, the Banach spaces $\ell^p(\mathbb{F})$, $p \in [1, \infty)$, have the property of being separable.

6.7.20 Proposition ($\ell^{\mathrm{p}}(\mathbb{F})$ is separable for p $\in [1, \infty)$) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $\mathrm{p} \in [1, \infty)$ then the Banach space $(\ell^{\mathrm{p}}(\mathbb{F}), \|\cdot\|_{\mathrm{p}})$ is separable.*

Proof We recall the definition of $\mathscr{D}_q(\mathbb{F})$ from the proof of Proposition 6.7.13 for $q \in \mathbb{Q}$. There we showed that $\mathscr{D}(\mathbb{F})$ was countable. We will show that $\mathscr{D}_0(\mathbb{R})$ is dense in $\ell^p(\mathbb{F})$. It is clear that $\mathscr{D}_0(\mathbb{F}) \subseteq \ell^p(\mathbb{F})$ for $p \in [1, \infty)$. To show that it is dense in $\ell^p(\mathbb{F})$ let $\epsilon \in \mathbb{R}_{>0}$ and let $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^p(\mathbb{F})$. Let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\Big(\sum_{j=N+1}^{\infty}|a_j|^p\Big)^{1/p} < \frac{\epsilon}{2}.$$

Now let $q_1, \dots, q_N \in \mathbb{Q}$ be such that

$$\Big(\sum_{j=1}^{N} |a_j - q_j|^p\Big)^{1/p} < \frac{\epsilon}{2}.$$

Then, taking $q_j = 0$ for $j > N$,

$$\big\|(a_j)_{j\in\mathbb{Z}_{>0}} - (q_j)_{j\in\mathbb{Z}_{>0}}\big\|_p = \Big(\sum_{j=1}^{N} |a_j - q_j|^p\Big)^{1/p} + \Big(\sum_{j=N+1}^{\infty} |a_j|^p\Big)^{1/p} < \epsilon.$$

Since $(q_j)_{j\in\mathbb{Z}_{>0}} \in \mathscr{D}_0(\mathbb{R})$ the result follows. ■

From this result we have the following useful corollary which finishes off Example 6.3.1–1.

6.7.21 Corollary ($\ell^p(\mathbb{F})$ is the completion of $\mathbb{F}_0^\infty$) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $p \in [1, \infty)$ then $(\ell^p(\mathbb{F}), \|\cdot\|_p)$ is the completion of $(\mathbb{F}_0^\infty, \|\cdot\|_p)$.*

Proof Borrowing the notation from the proof of Proposition 6.7.20 we have

$$\mathscr{D}_0(\mathbb{F}) \subseteq \mathbb{F}_0^\infty \subseteq \ell^p(\mathbb{F})$$

from which we deduce, using the proof of Proposition 6.7.20, that

$$\ell^p(\mathbb{F}) = \mathrm{cl}(\mathscr{D}_0(\mathbb{F})) \subseteq \mathrm{cl}(\mathbb{F}_0^\infty) \subseteq \mathrm{cl}(\ell^p(\mathbb{F})) = \ell^p(\mathbb{F}).$$

Therefore, $\mathrm{cl}(\mathbb{F}_0^\infty) = \ell^p(\mathbb{F})$, as desired. ■

6.7.3 Banach spaces of direct sums

Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $((\mathsf{V}_i, \|\cdot\|_i))_{i\in I}$ be a family of nontrivial $\mathbb{F}$-Banach spaces. We shall generalise the situation of Proposition 6.3.4 as follows. For $p \in [1, \infty]$ we define a norm $\|\cdot\|_{I,p}$ on $\bigoplus_{i\in I} \mathsf{V}_i$ by

$$\|(v_i)_{i\in I}\|_{I,p} = \begin{cases} \Big(\sum_{i\in I} \|v_i\|_i^p\Big)^{1/p}, & p \in [1, \infty), \\ \sup\{\|v_i\|_i \mid i \in I\}, & p = \infty. \end{cases}$$

The argument in the proof of Proposition 6.3.4 used to show incompleteness of $(\bigoplus_{i\in I} \mathsf{V}_i, \|\cdot\|_{I,1})$ when I is infinite is easily adapted to the case when $p \in [1, \infty)$. Moreover, for $p = \infty$ one can also show that $(\bigoplus_{i\in I} \mathsf{V}_i, \|\cdot\|_{I,\infty})$ is incomplete; we leave this to the reader as Exercise 6.3.6.

Note that the situation we consider here is a generalisation of the spaces of sequences considered in detail in Section 6.7.2. Indeed, the situation in Section 6.7.2 occurs upon taking $I = \mathbb{Z}_{>0}$ and $\mathsf{V}_i = \mathbb{F}$ for each $i \in I$. For this reason, many of the particulars in this section go just as they do in Section 6.7.2, and we encourage the reader to understand this. It will be helpful in understanding the further generalisations we will make from families to functions.

That $\|\cdot\|_p$ is a norm for each $p \in [1, \infty)$ is not difficult to show, but we will show this as we go along in any event. In fact, we shall follow closely the course set out in Section 6.7.2. In keeping with this, we start off making the following definition.

6.7.22 Definition ($\boldsymbol{\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i)}$) If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $((\mathsf{V}_i, \|\cdot\|_i)_{i\in I}$ is a family of normed $\mathbb{F}$-vector spaces then we define

$$\ell^\infty(\textstyle\bigoplus_{i\in I}\mathsf{V}_i) = \left\{(v_i)_{i\in I} \in \prod_{i\in I}\mathsf{V}_i \;\middle|\; \sup\{\|v_i\|_i \mid i \in I\} < \infty\right\}$$

and define

$$\|(v_i)_{i\in I}\|_{I,\infty} = \sup\{\|v_i\|_i \mid i \in I\}$$

for $(v_i)_{i\in I} \in \ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i$. •

It is evident (and see Exercise 6.7.5) that it is necessary that each of the normed vector spaces V_i be a Banach space if $\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i)$ is to be a Banach space. Moreover, this is sufficient.

6.7.23 Theorem ($\boldsymbol{(\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i), \|\cdot\|_{I,\infty})}$ is a Banach space) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $((\mathsf{V}_i, \|\cdot\|_i))_{i\in I}$ be a family of $\mathbb{F}$-Banach spaces. Then $(\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i), \|\cdot\|_{I,\infty})$ is an $\mathbb{F}$-Banach space.*

Proof The only not entirely trivial norm property to verify for $\|\cdot\|_{I,\infty}$ is the triangle inequality:

$$\begin{aligned}\|(u_i)_{i\in I} + (v_i)_{i\in I}\|_{I,\infty} &= \sup\{\|u_i + v_i\|_i \mid i \in I\}\\ &\le \sup\{\|u_i\|_i + \|v_i\|_i \mid i \in I\}\\ &= \sup\{\|u_i\|_i \mid i \in I\} + \sup\{\|v_i\|_i \mid i \in I\}\\ &= \|(u_i)_{i\in I}\|_\infty + \|(v_i)_{i\in I}\|_\infty,\end{aligned}$$

where we have used Proposition 2.2.27.

Now let us verify that $(\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i), \|\cdot\|_{I,\infty})$ is complete. We let $((v_i^{(l)})_{i\in I})_{l\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i)$. We claim that, for each $i \in I$, $(v_i^{(l)})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in V_i. To see this, let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left\|(v_i^{(l)})_{i\in I} - (v_i^{(k)})_{i\in I}\right\|_{I,\infty} < \epsilon$$

for $k, l \ge N$. Then, by definition of $\|\cdot\|_{I,\infty}$,

$$\left\|v_i^{(l)} - v_i^{(k)}\right\|_i < \epsilon$$

for $k, l \ge N$ and for $i \in I$. Thus $(v_i^{(l)})_{l\in\mathbb{Z}_{>0}}$ is indeed a Cauchy sequence, and so converges to some $v_i \in \mathsf{V}_i$. We now claim that the sequence $((v_i^{(l)})_{i\in I})_{l\in\mathbb{Z}_{>0}}$ converges to $(v_i)_{i\in I}$. To see this, let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that

$$\left\|(v_i^{(l)})_{i\in I} - (v_i^{(k)})_{i\in I}\right\|_{I,\infty} < \tfrac{\epsilon}{2}$$

for $k, l \ge N$. Thus

$$\left\|v_i^{(l)} - v_i^{(l)}\right\|_i < \tfrac{\epsilon}{2}, \qquad k, l \ge N.$$

Now, for fixed $i \in I$, let $N' \in \mathbb{Z}_{>0}$ be sufficiently large that $\left\|v_i^{(k)} - v_i\right\|_i < \frac{\epsilon}{2}$ for $k \ge N'$. In this case, if $l \ge N$ and $k \ge \max\{N, N'\}$, we have

$$\left\|v_i^{(l)} - v_i\right\|_i \le \left\|v_i^{(l)} - v_i^{(k)}\right\|_i + \left\|v_i^{(k)} - v_i\right\|_i < \epsilon.$$

Since this holds for each $i \in I$ we have

$$\left\|(v_i^{(l)})_{i\in I} - (v_i)_{i\in I}\right\|_{I,\infty} \le \epsilon,$$

as desired. ■

Again sticking with the plan of Section 6.7.2, let us consider a subspace of $\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i)$ that is analogous to the subspace $\mathsf{c}_0(\mathbb{F})$ of $\ell^\infty(\mathbb{F})$.

6.7.24 Definition ($\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$) If $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and if $((\mathsf{V}_i,\|\cdot\|_i)_{i\in I}$ is a family of normed $\mathbb{F}$-vector spaces then we define $\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$ to be the elements $(v_i)_{i\in I}\in\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i)$ with the property that, for each $\epsilon\in\mathbb{R}_{>0}$ the set $\{i\in I \mid \|v_i\|_i\geq\epsilon\}$ is finite. •

As with the corresponding conclusion in Section 6.7.2, we have the following result.

6.7.25 Theorem ($\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$ is a Banach space) *Let $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and let $((\mathsf{V}_i,\|\cdot\|_i))_{i\in I}$ be a family of $\mathbb{F}$-Banach spaces. Then $(\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i),\|\cdot\|_{I,\infty})$ is an $\mathbb{F}$-Banach space, and is moreover the completion of $\bigoplus_{i\in I}\mathsf{V}_i$ with respect to the norm $\|\cdot\|_{I,\infty}$.*

Proof Let $((v_i^{(l)})_{i\in I})_{l\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$. By Theorem 6.7.23 this means that the sequence converges to $(v_i)_{i\in I}\in\ell^\infty(\bigoplus_{i\in I}\mathsf{V}_i)$. We next show that $(v_i)_{i\in I}\in\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$. Let $\epsilon\in\mathbb{R}_{>0}$ and let $N\in\mathbb{Z}_{>0}$ be sufficiently large that

$$\left\|v_i-v_i^{(k)}\right\|_i<\tfrac{\epsilon}{2},\qquad k\geq N,\ i\in I.$$

For fixed $k\geq N$ let $J\subseteq I$ be a finite set such that $\left\|v_i^{(k)}\right\|<\frac{\epsilon}{2}$ for each $i\in I\setminus J$. Then, for $i\in I\setminus J$,

$$\|v_i\|_i\leq\left\|v_i-v_i^{(k)}\right\|_i+\left\|v_i^{(k)}\right\|_i<\epsilon,$$

which completes the proof that $\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$ is a Banach space.

To see that $\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$ is the completion of $\bigoplus_{i\in I}\mathsf{V}_i$, let $\epsilon\in\mathbb{R}_{>0}$ and let $(v_i)_{i\in I}\in\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$. Let $J\subseteq I$ be a finite set such that $\|v_i\|_i<\epsilon$ for each $i\in I\setminus J$. Then define $(u_i)_{i\in I}\in\bigoplus_{i\in I}\mathsf{V}_i$ by

$$u_i=\begin{cases}v_i, & i\in J,\\ 0_{\mathsf{V}_i}, & i\in I\setminus J.\end{cases}$$

It then follows immediately that $\|(v_i)_{i\in I}-(u_i)_{i\in I}\|_{I,p}<\epsilon$, and so $\bigoplus_{i\in I}\mathsf{V}_i$ is dense in $\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$. ■

Now let us turn to the case of $p\in[1,\infty)$.

6.7.26 Definition ($\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$) Let $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and let $((\mathsf{V}_i,\|\cdot\|_i))_{i\in I}$ be a family of normed $\mathbb{F}$-vector spaces. For $p\in[1,\infty)$ we define

$$\ell^p(\textstyle\bigoplus_{i\in I}\mathsf{V}_i)=\left\{(v_i)_{i\in I}\in\prod_{i\in I}\mathsf{V}_i\;\middle|\;\displaystyle\sum_{i\in I}\|v_i\|_i^p<\infty\right\}$$

and

$$\|(v_i)_{i\in I}\|_{I,p}=\Big(\sum_{i\in I}\|v_i\|_i^p\Big)^{1/p},$$

for $(v_i)_{i\in I}\in\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$. •

Since the sum in the definition of $\|\cdot\|_{I,p}$ for $p\in[1,\infty)$ is over a general index set, it must be interpreted as in Section 2.4.7 (see also Section 6.4.6).

We now have the expected result that $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$ is a Banach space.

6.7.27 Theorem ($(\ell^{\mathrm{p}}(\bigoplus_{\mathrm{i}\in\mathrm{I}}\mathsf{V}_{\mathrm{i}}), \|\cdot\|_{\mathrm{I,p}})$ is a Banach space) *Let $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and let $((\mathsf{V}_{\mathrm{i}},\|\cdot\|_{\mathrm{i}}))_{\mathrm{i}\in\mathrm{I}}$ be a family of $\mathbb{F}$-Banach spaces. Then $(\ell^{\mathrm{p}}(\bigoplus_{\mathrm{i}\in\mathrm{I}}\mathsf{V}_{\mathrm{i}}),\|\cdot\|_{\mathrm{I,p}})$ is an $\mathbb{F}$-Banach space, and is moreover the completion of $\bigoplus_{\mathrm{i}\in\mathrm{I}}\mathsf{V}_{\mathrm{i}}$ with respect to the norm $\|\cdot\|_{\mathrm{I,p}}$.*

Proof Let us first verify that $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$ is a subspace. We first consider the case of $p=1$. Let $(u_i)_{i\in I},(v_i)_{i\in I}\in\ell^1(\bigoplus_{i\in I}\mathsf{V}_i)$ and note that for each finite subset $J\subseteq I$ we have

$$\sum_{j\in J}\|u_j+v_j\|_j\le\sum_{j\in J}\|u_j\|_j+\sum_{j\in J}\|v_j\|_j\le\|(u_i)_{i\in I}\|_{I,1}+\|(v_i)_{i\in I}\|_{I,1},$$

where we have used the triangle inequality for $\|\cdot\|_i$, $i\in I$. Therefore, by definition of sums over arbitrary index sets,

$$\|(u_i)_{i\in I}+(v_i)_{i\in I}\|_{I,1}=\sum_{i\in I}\|u_i+v_i\|_i\le\|(u_i)_{i\in I}\|_{I,1}+\|(v_i)_{i\in I}\|_{I,1}. \tag{6.13}$$

This shows that $(u_i)_{i\in I}+(v_i)_{i\in I}\in\ell^1(\bigoplus_{i\in I}\mathsf{V}_i)$. If $\alpha\in\mathbb{F}$ we have

$$\sum_{i\in I}\|\alpha v_i\|_i=|\alpha|\sum_{i\in I}\|v_i\|_i$$

by Proposition 2.4.30 (noting that the sum is over a countable subset of I). Thus $\alpha(v_i)_{i\in I}\in\ell^1(\bigoplus_{i\in I}\mathsf{V}_i)$, which shows that $\ell^1(\bigoplus_{i\in I}\mathsf{V}_i)$ is a subspace of $\prod_{i\in I}\mathsf{V}_i$.

Now let $p\in(1,\infty)$ and let $(u_i)_{i\in I},(v_i)_{i\in I}\in\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$. For each finite subset $J\subseteq I$

$$\Big(\sum_{j\in J}\|u_j+v_j\|_j^p\Big)^{1/p}\le\Big(\sum_{j\in J}\|u_j\|_j^p\Big)^{1/p}+\Big(\sum_{j\in J}\|v_j\|)_j^p\Big)^{1/p}\le\|(u_i)_{i\in I}\|_{I,p}+\|(v_i)_{i\in I}\|_{I,p}$$

by Lemma 6.7.3. Therefore,

$$\|(u_i)_{i\in I}+(v_i)_{i\in I}\|_{I,p}=\Big(\sum_{i\in I}\|u_i+v_i\|_i^p\Big)^{1/p}\le\|(u_i)_{i\in I}\|_{I,p}+\|(v_i)_{i\in I}\|_{I,p}. \tag{6.14}$$

This shows that $(u_i)_{i\in I}+(v_i)_{i\in I}\in\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$. It is easy to see, just as for the case of $p=1$, that $\alpha(v_i)_{i\in I}\in\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$ if $\alpha\in\mathbb{F}$ and if $(v_i)_{i\in I}\in\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$, $p\in(1,\infty)$.

As we have shown the triangle inequality for $\|\cdot\|_{I,p}$ already in (6.13) and (6.14), and since the other norm properties for $\|\cdot\|_{I,p}$ hold trivially, it follows that $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$ is a normed vector space. It remains to show that it is complete. Let $((v_i^{(l)})_{i\in I})_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$. We claim that the sequence $(v_i^{(l)})_{l\in\mathbb{Z}_{>0}}$ is a Cauchy sequence for each $i\in I$. For every $\epsilon\in\mathbb{R}_{>0}$ there exists $N\in\mathbb{Z}_{>0}$ such that

$$\Big\|(v_i^{(k)})_{i\in I}-(v_i^{(l)})_{i\in I}\Big\|_{I,p}=\Big(\sum_{i\in I}\Big\|v_i^{(k)}-v_i^{(l)}\Big\|_i^p\Big)^{1/p}<\epsilon.$$

Now let $\epsilon\in\mathbb{R}_{>0}$ and let $N\in\mathbb{Z}_{>0}$ be such that

$$\Big\|(v_i^{(k)})_{i\in I}-(v_i^{(l)})_{i\in I}\Big\|_{I,p}<\epsilon.$$

Then

$$\Big\|v_i^{(k)}-v_i^{(l)}\Big\|_i^p\le\sum_{i\in I}\Big\|v_i^{(k)}-v_i^{(l)}\Big\|_i^p<\epsilon^p,$$

giving $(v_i^{(l)})_{l\in\mathbb{Z}_{>0}}$ as a Cauchy sequence. Denote its limit by $v_i \in \mathsf{V}_i$. We next claim that $(v_i^{(l)})_{l\in\mathbb{Z}_{>0}}$ converges to $(v_i)_{i\in I}$ in $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that

$$\left\|(v_i^{(l)})_{i\in I} - (v_i^{(k)})_{i\in I}\right\|_p < \tfrac{\epsilon}{2}, \qquad l,k \geq N.$$

For any finite subset $J \subseteq I$ we claim that the sequence $((v_j^{(l)})_{j\in J})_{l\in\mathbb{Z}}$ converges to $(v_j)_{j\in J}$ in $\bigoplus_{j\in J}\mathsf{V}_j$ with respect to the norm $\|\cdot\|_{J,p}$ defined by

$$\|(v_j)_{j\in J}\|_{J,p} = \Big(\sum_{j\in J}\|v_j\|_j^p\Big)^{1/p}.$$

This claim is proved for $p = 1$ in Proposition 6.3.4. The proof for $p \in (1,\infty)$ is exactly the same, save for notation. Thus there exists $N' \in \mathbb{Z}_{>0}$ such that

$$\Big(\sum_{j\in J}\left\|v_j^{(k)} - v_j\right\|_j^p\Big)^{1/p} < \frac{\epsilon}{2}, \qquad k \geq N'.$$

Then, for $k \geq \max\{N, N'\}$,

$$\begin{aligned}\Big(\sum_{j\in J}\left\|v_j^{(l)} - v_j\right\|_j^p\Big)^{1/p} &\leq \Big(\sum_{j\in J}\left\|v_j^{(l)} - v_j^{(k)}\right\|_j^p\Big)^{1/p} + \Big(\sum_{j\in J}\left\|v_j^{(k)} - v_j\right\|_j^p\Big)^{1/p} \\ &\leq \left\|(v_i^{(l)})_{i\in I} - (v_i^{(k)})_{i\in I}\right\|_{I,p} + \Big(\sum_{j\in J}\left\|v_j^{(k)} - v_j\right\|_j^p\Big)^{1/p} < \epsilon.\end{aligned}$$

Since this can be done for any finite set $J \subseteq I$ we have

$$\left\|(v_i^{(l)})_{i\in I} - (a_i)_{i\in I}\right\|_p \leq \epsilon.$$

This gives convergence of $(v_i^{(l)})_{l\in\mathbb{Z}_{>0}}$ to $(a_i)_{i\in I}$ in $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$, as desired. ■

Of significant interest is the case when I is finite. In this case, all of the Banach spaces $\ell^p(\bigoplus_{i\in I}\mathsf{V}_i)$, $p \in [1,\infty]$, and $\mathsf{c}_0(\bigoplus_{i\in I}\mathsf{V}_i)$ are the same and equal to $\bigoplus_{i\in I}\mathsf{V}_i$. In particular, $\bigoplus_{i\in I}\mathsf{V}_i$ is a Banach space if I is finite and if all of the normed vector spaces V_i, $i \in I$, are complete.

6.7.4 Banach spaces of continuous functions on $\mathbb{R}$

One way to think of this section is as giving a generalisation of the construction of $\ell^\infty(\mathbb{F})$ and its subspaces in Section 6.7.2. The generalisation is to functions on the real line from sequences, which can be thought of as functions on $\mathbb{Z}_{>0}$. For functions on the real line one has the possible property of continuity that one is compelled to keep track of.

We begin by providing the classes of continuous functions we will talk about. We recall from Definition **??** that if $I \subseteq \mathbb{R}$ is an interval and if $A \subseteq I$ then $\mathrm{cl}_I(A) = \mathrm{cl}(A) \cap I$.

6.7.28 Definition ($\mathsf{C}^0(I;\mathbb{F})$, $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$, $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$, $\mathsf{C}^0_0(I;\mathbb{F})$) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $I \subseteq \mathbb{R}$ be an interval.

(i) $\mathsf{C}^0(I;\mathbb{F}) = \{f : I \to \mathbb{F} \mid f \text{ is continuous}\}$.

(ii) If $f \in \mathsf{C}^0(I;\mathbb{F})$ then the ***support*** of f is*missing stuff*

$$\operatorname{supp}(f) = \operatorname{cl}_I(\{x \in I \mid f(x) \neq 0\}).$$

(iii) $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F}) = \{f \in \mathsf{C}^0(I;\mathbb{F}) \mid f \text{ has compact support}\}$.

(iv) $\mathsf{C}^0_0(I;\mathbb{F}) = \{f \in \mathsf{C}^0(I;\mathbb{F}) \mid$ for every $\epsilon \in \mathbb{R}_{>0}$ there exists a compact set $K \subseteq I$ such that $\{x \in I \mid |f(x)| \geq \epsilon\} \subseteq K\}$.

(v) $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F}) = \{f \in \mathsf{C}^0(I;\mathbb{F}) \mid$ there exists $M \in \mathbb{R}_{>0}$ such that $|f(x)| \leq M$ for all $x \in I\}$. •

One should be a little careful about the meaning of compact support when I is not closed. For example, the function $f \in \mathsf{C}^0_{\mathrm{bdd}}((0,1];\mathbb{F})$ defined by $f(x) = 1$ does not have compact support since its support is $(0,1]$.

We first understand the case when $I = \mathbb{R}$. In this case, one can verify that

$$\mathsf{C}^0_0(\mathbb{R};\mathbb{F}) = \Big\{f \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}) \;\Big|\; \lim_{|x|\to\infty} |f(x)| = 0\Big\} \tag{6.15}$$

(this is Exercise 6.7.6). Thus $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ consists of those functions which "die off" at infinity.

Clearly

$$\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F}) \subset \mathsf{C}^0_0(\mathbb{R};\mathbb{F}) \subset \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}) \subset \mathsf{C}^0(\mathbb{R};\mathbb{F}). \tag{6.16}$$

For $I = \mathbb{R}$ the vector space $\mathsf{C}^0(I;\mathbb{F})$ is too large to be of interest for the purposes of the discussion here. This is simply because continuous functions on $\mathbb{R}$ can be unbounded, and we wish to use a norm that is reliant on functions being bounded. Indeed, we define $\|\cdot\|_\infty$ by

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in \mathbb{R}\}$$

for $f \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$. That this is a norm follows just as do the norm properties of Example 6.1.3–10.

Let us get the ball rolling by giving an important property of $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$. This result should be thought of as being analogous to $(\mathbb{F}^\infty_0, \|\cdot\|_\infty)$ not being complete.

6.7.29 Proposition (($\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty$) is not complete) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *then* $(\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$ *is not complete.*

Proof Let us define a sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$ by

$$f_j(x) = \begin{cases} \frac{1}{1+x^2}, & x \in [-j, j], \\ 0, & \text{otherwise.} \end{cases}$$

Let $\epsilon \in \mathbb{R}_{>0}$. Since $\lim_{x\to\infty} \frac{1}{1+x^2} = 0$ it follows that there exists $N \in \mathbb{Z}_{>0}$ such that $\left|\frac{1}{1+x_1^2} - \frac{1}{1+x_2^2}\right| < \epsilon$ for every $x_1, x_2 \geq N$. It then holds that $|f_j(x) - f_k(x)| < \epsilon$ for every

$j, k \geq N$ and for every $x \in \mathbb{R}$. This shows that $(f_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. We next claim that this sequence does not converge. The argument used in the lemma in Example 6.3.1–2 can be adapted to show that if $g \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ is a function to which the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges then $g(x) = \frac{1}{1+x^2}$ for every $x \in \mathbb{R}$. In particular, the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ does not converge in $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$, and so $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$ is not complete. ■

With this in our back pocket let us proceed in a manner entirely analogous to what we did in Section 6.7.2 in looking at $\ell^\infty(\mathbb{F})$ and its subspaces. Here the key observation is the following fairly obvious translation from the language of Section 3.5.2 to the current language of convergence in normed vector spaces.

6.7.30 Proposition (Characterisation of convergence in $(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ is a sequence in $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ then the following statements are equivalent:*

(i) the sequence $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ converges uniformly to $\mathrm{f} \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$;

(ii) the sequence $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ converges to $\mathrm{f} \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ with respect to the norm $\|\cdot\|_\infty$.

Proof This just follows directly from the definitions of each sort of convergence. If the reader does not see this, they ought to convince themselves that this is the case. ■

The following theorem is now fairly easily proved, given what we already did in Section 3.5.2.

6.7.31 Theorem ($(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ then $(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$ is an $\mathbb{F}$-Banach space.*

Proof Let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$. By Theorem 3.5.8***missing stuff*** it follows that this sequence converges to a function $f \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$, and so the theorem follows. ■

As with $\ell^\infty(\mathbb{F})$, $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ is not separable.

6.7.32 Proposition ($\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ is not separable) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ then $(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$ is not separable.*

Proof Define a function $g_0 \colon \mathbb{R} \to \mathbb{F}$ by

$$g_0(x) = \begin{cases} 1 + x, & x \in [-\frac{1}{2}, 0], \\ 1 - x, & x \in (0, \frac{1}{2}], \\ 0, & \text{otherwise}. \end{cases}$$

Then let $\mathscr{U}$ be the collection of functions $f \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ of the form

$$f(x) = \sum_{j\in\mathbb{Z}_{>0}} (-1)^{k_j} g_0(x - j)$$

where $(k_j)_{j\in\mathbb{Z}_{>0}}$ is a sequence in $\{0, 1\}$. The reader ought to sketch the graph of a typical function in $\mathscr{U}$ to understand what they are doing. Upon doing this it will be clear that, if $f \in \mathscr{U}$ then $\|f\|_\infty = 1$ and if $f_1, f_2 \in \mathscr{U}$ are distinct then $\|f_1 - f_2\|_\infty = 2$. The remainder of the proof follows the proof of Proposition 6.7.10, but we give it here for completeness.

Note that there are as many distinct functions in $\mathscr{U}$ as there are maps from $\mathbb{Z}_{>0}$ into $\{0,1\}$. Thus $\operatorname{card}(\mathscr{U}) = 2^{\aleph_0}$. It then follows from Exercises **??**, **??**, and 2.1.4 that $\mathscr{U}$ is uncountable. By Exercise 6.1.3 we have

$$\begin{aligned} & |\|g\|_\infty - \|f\|_\infty| \le \|g-f\|_\infty \\ \implies\quad & |\|g\|_\infty - 1| \le 1 \\ \implies\quad & \left\|g\right\|_\infty \le 2 \end{aligned}$$

for $f \in \mathscr{U}$. Thus $\mathsf{B}(1,f) \subseteq \mathsf{B}(2, 0_{\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})})$ for each $f \in \mathscr{U}$. If $f, g \in \mathscr{U}$ are distinct, and $\alpha \in \mathsf{B}(1,f)$ and $\beta \in \mathsf{B}(1,g)$ then

$$\|\alpha - g\|_\infty \ge |\|\alpha - f\|_\infty - \|f-g\|_\infty| \ge 2$$

using Proposition **??**. Thus $\alpha \notin \mathsf{B}(1,g)$. One similarly shows that $\beta \notin \mathsf{B}(1,f)$. This shows that $\mathsf{B}(2, 0_{\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})})$ contains the collection

$$\{\mathsf{B}(1,f) \mid f \in \mathscr{U}\}$$

of disjoint open balls. In particular, if $(g_j)_{j\in\mathbb{Z}_{>0}}$ is any countable subset of $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ then there is a countable or finite subset $(f_\alpha)_{\alpha\in A}$ of $\mathscr{U}$ in which are contained all of the functions $(g_j)_{j\in\mathbb{Z}_{>0}}$. Note that

$$\operatorname{cl}((g_j)_{j\in\mathbb{Z}_{>0}}) \subseteq \cup_{\alpha\in A}\overline{\mathsf{B}}(1,f_\alpha).$$

Therefore, any of the set of balls

$$\{\mathsf{B}(1,f) \mid f \in \mathscr{U},\ f \neq f_\alpha,\ \alpha \in A\}$$

cannot lie in $\operatorname{cl}((g_j)_{j\in\mathbb{Z}_{>0}})$ which prohibits $(g_j)_{j\in\mathbb{Z}_{>0}}$ from being dense. ■

Next let us characterise the completion of $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$. The following result is entirely analogous to Corollary 6.7.14 which asserts that $\mathsf{c}_0(\mathbb{F})$ is the completion of $\mathbb{F}^\infty_0$.

6.7.33 Theorem ($(\mathsf{C}^0_0(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *then* $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ *is an* $\mathbb{F}$*-Banach space, and moreover is the completion of* $(\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$.

Proof We first make the observation that $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ is a subspace of $\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$. This follows from Propositions 2.3.23 and 2.3.29. Now suppose that $(f_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$. By Theorem 6.7.31 there exists a function $f \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$ such that $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to f. We need only show that $f \in \mathsf{C}^0_0(\mathbb{R};\mathbb{F})$. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be sufficiently large that $|f(x) - f_j(x)| < \frac{\epsilon}{2}$ for all $x \in \mathbb{R}$ provided that $j \ge N$. Let $K \subseteq \mathbb{R}$ be a compact set such that $|f_N(x)| < \frac{\epsilon}{2}$ for $x \in \mathbb{R} \setminus K$. Then, for $x \in \mathbb{R} \setminus K$ we have

$$|f(x)| \le |f(x) - f_N(x)| + |f_N(x)| < \epsilon,$$

giving $f \in \mathsf{C}^0_0(\mathbb{R};\mathbb{F})$, as desired.

To show that $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ is the completion of $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$, let $f \in \mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ and define $(f_j)_{j\in\mathbb{Z}_{>0}}$ by

$$f_j(x) = \begin{cases} f(x), & x \in [-j, j], \\ f(-j)(j+1+x), & x \in [-j-1, -j), \\ f(j)(j+1-x), & x \in (j, j+1], \\ 0, & \text{otherwise}. \end{cases}$$

We claim that this sequence converges to f. For $\epsilon \in \mathbb{R}_{>0}$ let $N \in \mathbb{Z}_{>0}$ have the property that $|f(x)| < \epsilon$ if $|x| \geq N$. Then we immediately have $|f(x) - f_j(x)| < \epsilon$ for $j \geq N$, giving the desired convergence, and showing that $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$ is dense in $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$. ■

Just as $\mathsf{c}_0(\mathbb{F})$ is separable, so too is $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$.

6.7.34 Proposition ($\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ is separable) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *then* $(\mathsf{C}^0_0(\mathbb{R};\mathbb{F}), \|\cdot\|_\infty)$ *is separable.*

Proof For $N \in \mathbb{Z}_{>0}$ let us denote by $P_N(\mathbb{F})$ the set of functions $f\colon \mathbb{R} \to \mathbb{F}$ having the form

$$f(x) = \begin{cases} z_k x^k + \dots + z_1 x + z_0, & x \in [-N, N], \\ (z_k N^k + \dots + z_1 N + z_0)(N + 1 - x), & x \in (N, N+1), \\ ((-1)^k z_k N^k + \dots - z_1 N + z_0)(N + 1 + x), & x \in (-N-1, -N), \\ 0, & |x| \geq N + 1, \end{cases}$$

where $k \in \mathbb{Z}_{\geq 0}$ and $z_0, z_1, \dots, z_k \in \mathbb{F}$ are rational if $\mathbb{F} = \mathbb{R}$ and whose real and imaginary parts are rational of $\mathbb{F} = \mathbb{C}$. Note that functions in $P_N(\mathbb{F})$ are continuous. Moreover, for each $N \in \mathbb{Z}_{>0}$ the set $P_N(\mathbb{F})$ is countable by Proposition **??**. Thus $\cup_{N \in \mathbb{Z}_{>0}} P_N(\mathbb{F})$ is also countable, again by Proposition **??**.

We claim that $\cup_{N \in \mathbb{Z}_{>0}} P_N(\mathbb{F})$ is dense in $\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$. Indeed, let $f \in \mathsf{C}^0_0(\mathbb{R};\mathbb{F})$ and let $\epsilon \in \mathbb{R}_{>0}$. Let $N \in \mathbb{Z}_{>0}$ be sufficiently large that $|f(x)| < \epsilon$ for $|x| \geq N$. By the Weierstrass Approximation Theorem, Theorem 3.5.21, let $g \in P_N(\mathbb{F})$ be such that $|f(x) - g(x)| < \epsilon$ for $x \in [-N, N]$. Our construction of functions in $P_N(\mathbb{F})$ then ensures that $|f(x) - g(x)| < \epsilon$ for all $x \in \mathbb{R}$. ■

In the preceding discussion we have pointed out various analogies with constructions concerning sequences in Section 6.7.2. In Table 6.1 we summarise the correspondences. The correspondences for the sequence spaces $\ell^p(\mathbb{F})$ for $p \in [1, \infty)$ are more complicated, and we present these in Table 6.2.

Table 6.1 The relationships between the objects in the left column are analogous to the relationships between the objects in the right column

Sequence space	Function space
$\mathbb{F}^\infty_0$	$\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$
$\ell^\infty(\mathbb{F})$	$\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F})$
$\mathsf{c}_0(\mathbb{F})$	$\mathsf{C}^0_0(\mathbb{R};\mathbb{F})$

Having now somewhat understood the structure of the spaces $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$, $\mathsf{C}^0_0(I;\mathbb{F})$, and $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$ when $I = \mathbb{R}$, let us turn to the case of a general interval. It is fairly easy to carry out the programme directly in this case, adapting the arguments above. However, it is also the case that we shall do this in some generality in Section 6.7.5. Therefore, we abbreviate the discussion somewhat, mostly only giving outlines of proofs and referring to the more general results for complete arguments.

First let us observe that Proposition 6.7.30 holds for arbitrary intervals.

6.7.35 Proposition (Characterisation of convergence in $(\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathbb{F}),\|\cdot\|_\infty)$) *If* $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$*, if* $\mathrm{I}\subseteq\mathbb{R}$*, and if* $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *is a sequence in* $\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F})$ *then the following statements are equivalent:*

(i) the sequence $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges uniformly to* $\mathrm{f}\in\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F})$*;*

(ii) the sequence $(\mathrm{f_j})_{\mathrm{j}\in\mathbb{Z}_{>0}}$ *converges to* $\mathrm{f}\in\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F})$ *with respect to the norm* $\|\cdot\|_\infty$*.*

Proof As with Proposition 6.7.30, this follows directly from the definitions. ■

Now let us indicate that things are significantly more trivial for compact intervals than for general intervals.

6.7.36 Proposition (Continuous function spaces for compact intervals) *If* $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ *and if* $\mathrm{I}\subseteq\mathbb{R}$ *is a compact interval, then*

$$\mathsf{C}^0_{\mathrm{cpt}}(\mathrm{I};\mathbb{F})=\mathsf{C}^0_0(\mathrm{I};\mathbb{F})=\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F})=\mathsf{C}^0(\mathrm{I};\mathbb{F}).$$

Proof This is a consequence of (6.16) along with the fact that $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})=\mathsf{C}^0(I;\mathbb{F})$ since every closed subset of I is compact according to Corollary 2.5.28. ■

For compact intervals this gives the following characterisation of their continuous functions as forming a particularly nice Banach space.

6.7.37 Corollary (Properties of continuous function spaces for compact intervals) *If* $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ *and if* $\mathrm{I}\subseteq\mathbb{R}$ *is a compact interval, then* $\mathsf{C}^0_{\mathrm{cpt}}(\mathrm{I};\mathbb{F})$*,* $\mathsf{C}^0_0(\mathrm{I};\mathbb{F})$*,* $\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F})$*, and* $\mathsf{C}^0(\mathrm{I};\mathbb{F})$ *are separable* $\mathbb{F}$*-Banach spaces with the norm* $\|\cdot\|_\infty$*.*

Proof That these are Banach spaces follows from Theorem 3.5.8***missing stuff*** since there we showed that in $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$ all Cauchy sequences converge. Separability follows from the Weierstrass Approximation Theorem, just as does Proposition 6.7.34. ■

Since $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ is the smallest of the spaces we consider, let us characterise precisely when it is a Banach space.

6.7.38 Proposition (Completeness of $(\mathsf{C}^0_{\mathrm{cpt}}(\mathrm{I};\mathbb{F}),\|\cdot\|_\infty)$) *If* $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ *and if* $\mathrm{I}\subseteq\mathbb{R}$ *is an interval, then* $(\mathsf{C}^0_{\mathrm{cpt}}(\mathrm{I};\mathbb{F}),\|\cdot\|_\infty)$ *is complete if and only if* I *is compact.*

Proof For the noncompleteness of $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ when I is not compact, we consider two cases of intervals: $I=(0,1]$ and $I=[0,\infty)$. The proof for an arbitrary noncompact interval follows by a trivial modification of these two cases.

First we show that $\mathsf{C}^0_{\mathrm{cpt}}((0,1];\mathbb{F})$ is not complete. We consider a sequence of functions $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{C}^0_{\mathrm{cpt}}((0,1];\mathbb{F})$ defined by

$$f_j(x)=\begin{cases}0, & x\in(0,\frac{1}{j}],\\ 2\frac{jx-1}{j-2}, & x\in[\frac{1}{j},\frac{1}{2}],\\ 1, & x\in[\frac{1}{2},1].\end{cases}$$

The reader is encouraged to plot the graphs of a few of the functions in this sequence to see what they are doing. Upon doing this it is easy to see that the sequence converges pointwise, in fact uniformly, to the function $f\colon(0,1]\to\mathbb{F}$ defined by

$$f(x)=\begin{cases}x, & x\in(0,\frac{1}{2}],\\ 1, & x\in(\frac{1}{2},1].\end{cases}$$

We leave the elementary formal verification of this to the reader. Thus the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges in the normed vector space $(\mathsf{C}^0_{\mathrm{bdd}}((0,1];\mathbb{F}),\|\cdot\|_\infty)$. It is, therefore, a Cauchy sequence. However, since f does not have compact support, the sequence does not converge in $\mathsf{C}^0_{\mathrm{cpt}}((0,1];\mathbb{F})$, giving the incompleteness of $\mathsf{C}^0_{\mathrm{cpt}}((0,1];\mathbb{F})$.

Now we show that $\mathsf{C}^0_{\mathrm{cpt}}([0,\infty);\mathbb{F})$ is not complete. Let us define a sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{C}^0_{\mathrm{cpt}}([0,\infty);\mathbb{F})$ by

$$f_j(x)=\begin{cases}\frac{1}{1+x^2}, & x\in[0,j],\\ 0, & \text{otherwise.}\end{cases}$$

It then follows, just as in the proof of Proposition 6.7.29, that this is a Cauchy sequence that does not converge.

That $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ is complete when I is compact is Proposition 6.7.36. ■

The bounded continuous functions on I form a Banach space.

6.7.39 Theorem ($(\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F}),\|\cdot\|_\infty)$ is a Banach space) *If $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and if $\mathrm{I}\subseteq\mathbb{R}$ is an interval then $(\mathsf{C}^0_{\mathrm{bdd}}(\mathrm{I};\mathbb{F}),\|\cdot\|_\infty)$ is an $\mathbb{F}$-Banach space. This Banach space is separable if and only if I is compact.*

Proof While the first assertion follows from Theorem 3.5.8***missing stuff*** just as does Theorem 6.7.31, we give a complete self-contained proof here, since this is an important result for us.

Let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a Cauchy sequence in $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$ and for $x\in I$ define $f(x)=\lim_{j\to\infty}f_j(x)$. This pointwise limit exists since $(f_j(x))_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{R}$ (why?).

First we claim that for any $\epsilon>0$ there exists $N\in\mathbb{Z}_{>0}$ such that $|f(x)-f_j(x)|<\epsilon$ for all $x\in I$ whenever $j\geq N$. Let $\epsilon\in\mathbb{R}_{>0}$ and let $x\in I$. Since $(f_j)_{j\in\mathbb{Z}_{>0}}$ is Cauchy there exists $N\in\mathbb{Z}_{>0}$ such that $|f_j(x)-f_k(x)|<\frac{\epsilon}{2}$. We may also find $N(x)\in\mathbb{Z}_{>0}$ such that $|f(x)-f_j(x)|<\frac{\epsilon}{2}$ for $j\geq N(x)$. Let $k=\max\{N,N(x)\}$. For $j\geq N$ we then have

$$\begin{aligned}|f_j(x)-f(x)|&=|(f_j(x)-f_k(x))+(f_k(x)-f(x))|\\&\leq|f_j(x)-f_k(x)|+|f_k(x)-f(x)|\\&<\tfrac{\epsilon}{2}+\tfrac{\epsilon}{2}=\epsilon,\end{aligned}$$

where we have used the triangle inequality. Note that this shows uniform convergence to f of the sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$, and so convergence to f using the norm $\|\cdot\|_\infty$.

We next claim that f is bounded. To see this, for $\epsilon>0$ let $N\in\mathbb{Z}_{>0}$ have the property that $\|f-f_N\|_\infty<\epsilon$. Then

$$|f(x)|\leq|f(x)-f_N(x)|+|f_N(x)|\leq\epsilon+\|f_N\|_\infty.$$

Since the expression on the right is independent of x, this gives the desired boundedness of f.

Finally we prove that the limit function f is continuous. As we showed above, for any $\epsilon>0$ there exists $N\in\mathbb{Z}_{>0}$ such that $|f_N(x)-f(x)|<\frac{\epsilon}{3}$ for all $x\in I$. Now fix $x_0\in I$, and consider the $N\in\mathbb{Z}_{>0}$ just defined. By continuity of f_N, there exists $\delta>0$ such that if $x\in I$ satisfies $|x-x_0|<\delta$, then $|f_N(x)-f_N(x_0)|<\frac{\epsilon}{3}$. Then, for $x\in I$ satisfying

$|x - x_0| < \delta$, we have

$$\begin{aligned}|f(x) - f(x_0)| &= |(f(x) - f_N(x)) + (f_N(x) - f_N(x_0)) + (f_N(x_0) - f(x_0))| \\ &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon,\end{aligned}$$

where we have again used the triangle inequality. Since this argument is valid for any $x_0 \in I$, it follows that f is continuous.

Now let us turn to the separability of $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$. The separability of $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$ when I is compact is part of Corollary 6.7.37. If I is not compact, there are two cases to consider, when I is bounded and when I is not bounded. If I is not bounded a modification of the argument used in Proposition 6.7.32 can be used to show that $\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F})$ is not separable. Thus we need only consider the case when I is bounded but not compact.

We consider the case of $I = (0,1]$, the general case following, *mutatis mutandis*, from this. For $j \in \mathbb{Z}_{>0}$ define $g_j\colon (0,1] \to \mathbb{F}$ by

$$g_j(x) = \begin{cases} 2j(Herex(j+1) - 1), & x \in [\frac{1}{j+1}, \frac{1+2j}{2j(j+1)}], \\ 2(j+1)(1 - jx), & x \in (\frac{1+2j}{2j(j+1)}, \frac{1}{j}], \\ 0, & \text{otherwise}. \end{cases}$$

The reader would probably benefit from sketching the graph of this function to understand what the proof is achieving. We now let $\mathscr{U}$ be the collection of functions $f \in \mathsf{C}^0_{\mathrm{bdd}}((0,1];\mathbb{F})$ of the form

$$f(x) = \sum_{j\in\mathbb{Z}_{>0}} (-1)^{k_j} g_j(x).$$

One can now repeat the argument of Proposition 6.7.32 using this collection $\mathscr{U}$ of functions to show that $\mathsf{C}^0_{\mathrm{bdd}}((0,1];\mathbb{F})$ is not separable. ■

The generalisation of Theorem 6.7.33 also holds.

6.7.40 Theorem ((C$^0_0$(I; $\mathbb{F}$), ||·||$_\infty$) is a Banach space) *If* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ *and if* $\mathrm{I} \subseteq \mathbb{R}$ *is an interval then* $(\mathsf{C}^0_0(\mathrm{I};\mathbb{F}), \|\cdot\|_\infty)$ *is a separable* $\mathbb{F}$*-Banach space, and moreover, is the completion of* $(\mathsf{C}^0_{\mathrm{cpt}}(\mathrm{I};\mathbb{F}), \|\cdot\|_\infty)$.

Proof A modification of the proof of Theorem 6.7.33 is easily made to give a direct proof; we leave the details to the reader. We also note that the present theorem also follows directly from the more general Theorem 6.7.43 below. ■

6.7.5 Banach spaces of continuous functions on metric spaces

We let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let (S,d) be a metric space and define

$$\mathsf{C}^0_{\mathrm{bdd}}(S;\mathbb{F}) = \{f\colon S \to \mathbb{F} \mid f \text{ is continuous and bounded}\}.$$

For $f \in \mathsf{C}^0_{\mathrm{bdd}}(S;\mathbb{F})$ we define

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in S\}.$$

We claim that $(\mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space.

6.7.41 Theorem ($(\mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F}),\|\cdot\|_\infty)$ is a Banach space) *$(\mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F}),\|\cdot\|_\infty)$ is a Banach space.*

Proof ***missing stuff*** First let us show that $\|\cdot\|_\infty$ is a norm. It is clear that $\|\lambda f\|_\infty = |\lambda|\|f\|_\infty$ for all $\lambda \in \mathbb{F}$ and $f \in \mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F})$, and that $\|f\|_\infty \geq 0$ and $\|f\|_\infty = 0$ if and only if $f = 0$. We also compute, using Proposition 2.2.27,

$$\begin{aligned}\|f+g\|_\infty &= \sup\{|f(x)+g(x)| \mid x \in S\}\\ &\leq \sup\{|f(x)+g(y)| \mid (x,y) \in S \times S\}\\ &\leq \sup\{|f(x)| \mid x \in S\} + \sup\{|g(y)| \mid y \in S\}\\ &= \|f\|_\infty + \|g\|_\infty\end{aligned}$$

for $f, g \in \mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F})$. To show that $(\mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F}),\|\cdot\|_\infty)$ is a complete normed vector space, we note that the norm topology is exactly the metric topology defined in general in Theorem **??**. Since $(\mathbb{F},|\cdot|)$ is complete, it then follows from Theorem **??** that $(\mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F}),\|\cdot\|_\infty)$ is also complete. ■

Let us record some of the properties of the Banach space $\mathsf{C}^0_{\mathrm{bdd}}(S,\mathbb{F})$.

6.7.42 Proposition (Properties of $\mathsf{C}^0_{\mathrm{bdd}}(S;\mathbb{F})$) ***missing stuff***

6.7.43 Theorem ($(\mathsf{C}^0_0(S;\mathbb{F}),\|\cdot\|_\infty)$ is a Banach space)

Proof ■

6.7.6 Banach spaces of continuous functions on locally compact topological spaces

6.7.7 Banach spaces of integrable functions on $\mathbb{R}$

In this section we look at an extremely important class of Banach spaces. In some sense, these are adaptations of the spaces of sequences considered in Section 6.7.2 to functions defined on intervals. These classes of functions play an essential rôle in Fourier analysis as we shall see in Chapters 12 and 13.

We begin, as we did with sequences, by considering functions that are, in the appropriate sense, bounded.

6.7.44 Definition ($\mathsf{L}^{(\infty)}(I;\mathbb{F})$) Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $I \subseteq \mathbb{R}$ be an interval. A measurable function $f\colon I \to \mathbb{F}$ is ***essentially bounded*** if there exists $M \in \mathbb{R}_{\geq 0}$ such that the set

$$\lambda(\{x \in I \mid |f(x)| > M\}) = 0.$$

The set of essentially bounded functions from I to $\mathbb{F}$ is denoted by $\mathsf{L}^{(\infty)}(I;\mathbb{F})$ and define

$$\|f\|_\infty = \inf\{M \in \mathbb{R}_{\geq 0} \mid \lambda(\{x \in I \mid |f(x)| > M\}) = 0\}$$

for $f \in \mathsf{L}^{(\infty)}(I;\mathbb{F})$. •

Let us give some initial properties of $\mathsf{L}^{(\infty)}(I;\mathbb{F})$.

6.7.45 Proposition (Properties of $(\mathsf{L}^{(\infty)}(\mathrm{I};\mathbb{F}),\|\cdot\|_\infty)$) *If* $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ *and if* $\mathrm{I}\subseteq\mathbb{R}$ *is an interval then* $(\mathsf{L}^{(\infty)}(\mathrm{I};\mathbb{F}),\|\cdot\|_\infty)$ *is a seminormed* $\mathbb{F}$*-vector space. Moreover,* $\|\mathrm{f}\|_\infty=0$ *if and only if* $\mathrm{f}(\mathrm{x})=0$ *for almost every* $\mathrm{x}\in\mathrm{I}$.

Proof The only seminorm property that is not completely trivial is the triangle inequality, so let us verify this. If $f\colon\phi\to\mathbb{R}$ is an arbitrary measurable function we denote***missing stuff***

$$\operatorname{ess\,sup}\{\phi(x)\mid x\in I\}=\inf\{M\in\mathbb{R}_{\geq 0}\mid \lambda(\{x\in I\mid \phi(x)>M\}=0)\}.$$

If

$$Z_\phi=\{x\in I\mid \phi(x)>\operatorname{ess\,sup}\{\phi(x)\mid x\in I\}\}$$

and if Z is any set of measure zero containing Z_ϕ then

$$\operatorname{ess\,sup}\{\phi(x)\mid x\in I\}=\sup\{\phi(x)\mid x\in I\setminus Z\}.$$

Now let $f,g\in\mathsf{L}^{(\infty)}(I;\mathbb{F})$ and compute

$$\begin{aligned}\|f+g\|_\infty &= \operatorname{ess\,sup}\{|f(x)+g(x)|\mid x\in I\}\\ &= \sup\{|f(x)+g(x)|\mid x\in I\setminus Z_{|f+g|}\}\\ &\leq \sup\{|f(x)|+|g(x)|\mid x\in I\setminus Z_{|f+g|}\}\\ &\leq \sup\{|f(x)|\mid x\in I\setminus Z_{|f+g|}\}+\sup\{|f(x)|\mid x\in I\setminus Z_{|f+g|}\}\\ &= \sup\{|f(x)|\mid x\in I\setminus(Z_{|f+g|}\cup Z_f)\}+\sup\{|f(x)|\mid x\in I\setminus(Z_{|f+g|}\cup Z_g)\}\\ &\leq \sup\{|f(x)|\mid x\in I\setminus Z_f\}+\sup\{|f(x)|\mid x\in I\setminus Z_g\}\\ &= \|f\|_\infty+\|g\|_\infty.\end{aligned}$$

Thus $(\mathsf{L}^{(\infty)}(I;\mathbb{F}),\|\cdot\|_\infty)$ is a seminormed $\mathbb{F}$-vector space, as claimed.

The final assertion of the result is clear. ■

Now let

$$Z^\infty(I;\mathbb{F})=\{f\in\mathsf{L}^{(\infty)}(I;\mathbb{F})\mid \|f\|_\infty=0\}.$$

By Theorem 6.1.8 we know that $\mathsf{L}^{(\infty)}(I;\mathbb{F})/Z^\infty(I;\mathbb{F})$,—i.e., the set of equivalence classes in $\mathsf{L}^{(\infty)}(I;\mathbb{F})$ where functions are equivalent if they agree almost everywhere—is a normed $\mathbb{F}$-vector space where the norm on the equivalence class $f+Z^\infty(I;\mathbb{F})$ is defined by

$$\|f+Z^\infty(I;\mathbb{F})\|_\infty=\|f\|_\infty;$$

it is convenient to use the same symbol for the norm.

6.7.46 Definition ($\mathsf{L}^\infty(\mathrm{I};\mathbb{F})$) For $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ and for an interval $I\subseteq\mathbb{R}$,

$$\mathsf{L}^\infty(I;\mathbb{F})=\mathsf{L}^{(\infty)}(I;\mathbb{F})/Z^\infty(I;\mathbb{F}).$$ •

Let us verify that $\mathsf{L}^\infty(I;\mathbb{F})$ is a Banach space.

6.7.47 Theorem ($(\mathsf{L}^\infty(\mathrm{I};\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $\mathrm{I} \subseteq \mathbb{R}$ then $(\mathsf{L}^\infty(\mathrm{I};\mathbb{F}), \|\cdot\|_\infty)$ is an $\mathbb{F}$-Banach space.*

Proof For brevity, let us denote $[f] = f + Z^\infty(I;\mathbb{F})$ the equivalence class of $f \in \mathsf{L}^{(\infty)}(I;\mathbb{F})$ in $\mathsf{L}^\infty(I;\mathbb{F})$. We use the characterisation of completeness of Theorem 6.4.6. We let $\sum_{j=1}^\infty [f_j]$ be an absolutely convergent series. For $j \in \mathbb{Z}_{>0}$ define

$$Z_j = \{x \in I \mid |f_j(x)| > \|f_j\|_\infty\},$$

noting that $\lambda(Z_j) = 0$. For $x \notin \cup_{j=1}^\infty Z_j$ we have

$$\sum_{j=1}^\infty |f_j(x)| \le \sum_{j=1}^\infty \|f_j\|_\infty = \sum_{j=1}^\infty \|[f_j]\|_\infty < \infty$$

since $\sum_{j=1}^\infty [f_j]$ is absolutely convergent. This means that $\sum_{j=1}^\infty f_j(x)$ converges since absolute convergence in $\mathbb{F}$ implies convergence by Proposition 2.4.3.***missing stuff*** Now define

$$f(x) = \begin{cases} \sum_{j=1}^\infty f_j(x), & x \notin \cup_{j=1}^\infty Z_j \\ 0, & \text{otherwise.} \end{cases}$$

By Proposition 5.6.18 the function f is measurable. We then have

$$\begin{aligned} & f(x) - \sum_{j=1}^n f_j(x) = \sum_{j=n+1}^\infty f_j(x), \quad x \notin \cup_{j=1}^\infty Z_j \\ \implies \quad & \Big\| f - \sum_{j=1}^n f_j \Big\|_\infty \le \sum_{j=n+1}^\infty \|f_j\|_\infty \\ \implies \quad & \Big\| \Big[f - \sum_{j=1}^n f_j \Big] \Big\|_\infty \le \sum_{j=n+1}^\infty \|[f_j]\|_\infty \\ \implies \quad & \lim_{n\to\infty} \Big\| [f] - \sum_{j=1}^n [f_j] \Big\|_\infty \le \lim_{n\to\infty} \sum_{j=n+1}^\infty \|[f_j]\|_\infty = 0, \end{aligned}$$

thus giving convergence of $\sum_{j=1}^\infty [f_j]$ to $[f]$ in $\mathsf{L}^\infty(I;\mathbb{F})$. ■

6.7.48 Notation (Representing functions in $\mathsf{L}^\infty(\mathrm{I};\mathbb{F})$) While functions in $\mathsf{L}^\infty(I;\mathbb{F})$ are, by definition, equivalence classes of functions in $\mathsf{L}^{(\infty)}(I;\mathbb{F})$. The usual convention, however, is to in practice identify the equivalence class with one of its representatives. Most of the time the identification of an equivalence class with one of its representatives does not cause problems. However, there do arise instances where the distinction between these things becomes important, and so one must keep in mind what one is actually doing in writing "f" rather than "$f + Z^\infty(I;\mathbb{F})$." •

As with its brother $\ell^\infty(\mathbb{F})$, $\mathsf{L}^\infty(I;\mathbb{F})$ is not separable.

6.7.49 Proposition ($\mathsf{L}^\infty(\mathrm{I};\mathbb{F})$ is not separable) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $\mathrm{I} \subseteq \mathbb{R}$ is an interval with a nonempty interior then $\mathsf{L}^\infty(\mathrm{I};\mathbb{F})$ is not separable.*

Proof We shall only sketch the argument here as the details are already present in the proof of Proposition 6.7.32 and Theorem 6.7.39. If I is not bounded then an appropriate adaptation of the proof of the proof of Proposition 6.7.32 can be used to show that $\mathsf{L}^\infty(I;\mathbb{F})$ is not separable. If I is bounded then the idea in the proof of Theorem 6.7.39 can be used to give non-separability of $\mathsf{L}^\infty(I;\mathbb{F})$ in this case. Note that functions in $\mathsf{L}^\infty(I;\mathbb{F})$ are not required to be continuous and so the idea in the proof of Theorem 6.7.39 does indeed carry over to all bounded intervals, even those that are compact. ■

Before we leave $\mathsf{L}^\infty(I;\mathbb{F})$ to talk about the spaces $\mathsf{L}^p(I;\mathbb{F})$ for $p \in [1,\infty)$ let us point out a possible source of confusion. We note that the Banach space $(\mathsf{L}^\infty(I;\mathbb{F}), \|(\|\cdot)_\infty)$ contains the Banach spaces $(\mathsf{C}^0_{\mathrm{bdd}}(I;\mathbb{F}); \|\cdot\|_\infty)$ and $\mathsf{C}^0_0(I;\mathbb{F})$ as a closed proper subspaces (they is a closed by Proposition 6.6.16 since it is complete). Thus $\mathsf{L}^\infty(I;\mathbb{F})$ is *not* the completion of these spaces. This is to be contrasted with the conclusion of Theorem 6.7.56 where we show that $\mathsf{L}^p(I;\mathbb{F})$ *is* the completion of a space of continuous functions when $p \in [1,\infty)$. This explains why the reader does not see $\mathsf{L}^\infty(I;\mathbb{F})$ in Tables 6.1 and 6.2.

Next we consider functions defined by their integrals. This is analogous to the sequence spaces $\ell^p(\mathbb{F})$, $p \in [1,\infty)$, being defined by their infinite sums.

6.7.50 Definition ($\mathsf{L}^{(p)}(\mathrm{I};\mathbb{F})$) Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$, let $p \in [1,\infty)$, and let $I \subseteq \mathbb{R}$ be an interval. Define a subspace $\mathsf{L}^{(p)}(I;\mathbb{F})$ of the measurable functions from I to $\mathbb{F}$ by

$$\mathsf{L}^{(p)}(I;\mathbb{F}) = \Big\{ f\colon I \to \mathbb{F} \;\Big|\; f \text{ measurable}, \int_I |f|^p \,\mathrm{d}\lambda < \infty \Big\}$$

and define

$$\|f\|_p = \Big(\int_I |f|^p \,\mathrm{d}\lambda\Big)^{1/p}$$

for $f \in \mathsf{L}^{(p)}(I;\mathbb{F})$. •

In the preceding definition it turns out to be crucial that the integral used is the Lebesgue integral. Indeed, many of the results we prove in this section simply do not hold if we instead attempt to use the Riemann integral. We shall, nonetheless, generally adopt the policy of writing the Lebesgue integral as $\int \mathrm{d}x$ rather than $\int \mathrm{d}\lambda$ for simplicity.

Let us give the analogues of Lemmata 6.7.16 and 6.7.17 in this setup.

6.7.51 Lemma (Hölder's inequality) *Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $\mathrm{I} \subseteq \mathbb{R}$ be an interval, and let $\mathrm{p} \in (1,\infty)$ with p' defined by $\frac{1}{\mathrm{p}} + \frac{1}{\mathrm{p}'} = 1$. Then, for $\mathrm{f} \in \mathsf{L}^{(\mathrm{p})}(\mathrm{I};\mathbb{F})$ and $\mathrm{g} \in \mathsf{L}^{(\mathrm{p}')}(\mathrm{I};\mathbb{F})$, $\mathrm{fg} \in \mathsf{L}^{(1)}(\mathrm{I};\mathbb{F})$ and*

$$\|\mathrm{fg}\|_1 \leq \|\mathrm{f}\|_{\mathrm{p}} \|\mathrm{g}\|_{\mathrm{p}'}.$$

Moreover, equality holds if and only if there exists $\alpha, \beta \in \mathbb{R}_{\geq 0}$, not both zero, such that

$$\alpha|\mathrm{f}(\mathrm{x})|^{\mathrm{p}} = \beta|\mathrm{g}(\mathrm{x})|^{\mathrm{p}'}, \qquad \textit{a.e. } \mathrm{x} \in \mathrm{I}.$$

Proof For $p, p' \in (1, \infty)$ satisfying $\frac{1}{p} + \frac{1}{p'} = 1$ we claim that for $x, y \in \mathbb{R}_{\geq 0}$ we have

$$xy \leq \frac{x^p}{p} + \frac{y^{p'}}{p'}.$$

This is trivial if either x or y are zero. So suppose that $x, y \in \mathbb{R}_{>0}$. Taking $\xi = \frac{x^p}{y^{p'}}$ we easily check that

$$xy \leq \frac{x^p}{p} + \frac{y^{p'}}{p'} \quad \Longleftrightarrow \quad \xi^{1/p} \leq \frac{\xi}{p} + \frac{1}{p'}.$$

One can check using Theorem 3.2.16 that the function

$$\xi \mapsto \frac{\xi}{p} + \frac{1}{p'} - \xi^{1/p}$$

has a minimum value of 0 attained at $\xi = 1$. Thus

$$\frac{\xi}{p} + \frac{1}{p'} - \xi^{1/p} \geq 0 \implies \qquad \implies xy \leq \frac{x^p}{p} + \frac{y^{p'}}{p'},$$

as desired.

Now let us proceed with the proof. The result is clearly true if $\|f\|_p = 0$ or $\|g\|_{p'} = 0$. So we assume neither of these are true. For all $x \in I$ we have

$$|f(x)g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^{p'}}{p'}.$$

Therefore, if $\|f\|_p = \|g\|_{p'} = 1$, we immediately have

$$\|fg\|_1 \leq \frac{1}{p} + \frac{1}{p'} = \|f\|_p \|g\|_{p'}.$$

In general we have

$$\|fg\|_1 = \|f\|_p \|g\|_{p'} \left\| \frac{f}{\|f\|_p} \frac{g}{\|g\|_{p'}} \right\|_1 \leq 1,$$

and the first part of the result follows.

If one chases through the argument above one sees that equality is achieved only when

$$|f(x)g(x)| \frac{|f(x)|^p}{p} + \frac{|g(x)|^{p'}}{p'}$$

for almost every $x \in I$. A tedious argument like that for the last part of Lemma 6.7.1, but replacing sums with integrals, shows that the above equality implies the final conclusion of the lemma. ■

There is a version of Hölder's inequality for the case when $p = 1$, and we refer to Exercise 6.7.8 for this.

Let us prove the Minkowski inequality in this case.

6.7.52 Lemma (Minkowski's inequality) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $I \subseteq \mathbb{R}$ *be an interval, and let* $p \in [1, \infty)$*. Then, for* $f, g \in \mathsf{L}^{(p)}(I; \mathbb{F})$*, we have* $f + g \in \mathsf{L}^{(p)}(I; \mathbb{F})$ *and*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Moreover, equality holds if and only if the following conditions hold:

(i) $p = 1$*: there exists nonnegative measurable functions* $\alpha, \beta\colon I \to \mathbb{R}_{\geq 0}$ *such that* $\alpha(x)f(x) = \beta(x)g(x)$ *and* $\alpha(x)$ *and* $\beta(x)$ *are not both zero for almost every* $\in I$*;*

(ii) $p \in (1, \infty)$*: there exists* $\alpha, \beta \in \mathbb{R}_{\geq 0}$*, not both zero, such that* $\alpha f(x) = \beta g(x)$ *for almost every* $x \in I$*.*

Proof For $p = 1$ we have

$$\|f + g\|_1 = \int_I |f(x) + g(x)| \, dx \leq \int_I |f(x)| \, dx + \int_I |g(x)| \, dx = \|f\|_1 + \|g\|_1.$$

The second assertion of the lemma for $p = 1$ follows from the fact, pointed out in the proof of Lemma 6.7.1, that $|a+b| = |a|+|b|$ for $a, b \in \mathbb{F}$ if and only if $\alpha a = \beta b$ for $\alpha, \beta \in \mathbb{R}_{\geq 0}$ not both zero. Note that the sets

$$A_f = \{x \in I \mid f(x) = 0\}, \quad A_g = \{x \in I \mid g(x) = 0\}, \quad A_{f,g} = \{x \in I \mid f(x)g(x) = 0\}$$

are measurable and so, therefore, are their complements. We then define $\alpha, \beta\colon I \to \mathbb{R}_{\geq 0}$ by

$$\alpha(x) = \begin{cases} g(x), & x \in I \setminus A_{f,g}, \\ g(x), & x \in A_{f,g} - A_f, \\ 0, & x \in A_g \end{cases}$$

and

$$\beta(x) = \begin{cases} f(x), & x \in I \setminus A_{f,g}, \\ 0, & x \in A_g, \\ f(x), & x \in A_{f,g} - A_g. \end{cases}$$

For $p \in (1, \infty)$ we let $\frac{1}{p} + \frac{1}{p'} = 1$. We then have

$$\left(|f(x) + g(x)|^{p-1}\right)^{p'} = |f(x) + g(x)|^p$$

from which we deduce that $|f + g|^{p-1} \in \mathsf{L}^{(p')}(I; \mathbb{F})$. Therefore, using Lemma 6.7.51,

$$\begin{aligned} \int_I |f(x) + g(x)|^p \, dx &\leq \int_I |f(x)||f(x) + g(x)|^{p-1} \, dx + \int_I |g(x)||f(x) + g(x)|^{p-1} \, dx \\ &\leq \|f\|_p \||f + g|^{p-1}\|_{p'} + \|g\|_p \||f + g|^{p-1}\|_{p'} \\ &= (\|f\|_p + \|g\|_p)\Big(\int_I |f(x) + g(x)|^p \, dx\Big)^{1/p'}, \end{aligned}$$

which implies that

$$\|f + g\|_p^{p - p/p'} \leq \|f\|_p + \|g\|_p,$$

provided that $\|f + g\|_p \neq 0$ (if it is zero, the result is trivial). The first part of the result follows since $p - p/p' = 1$. The second part of the result for $p \in (1, \infty)$ follows as does the second part of the proof of Lemma 6.7.1, replacing "for every $j \in \{1, \ldots, n\}$" with "for almost every $x \in I$" and replacing "for some $j \in \{1, \ldots, n\}$" with "for $x \in A$ with $A \subseteq I$ of positive measure." We leave the tedious details to the reader. ■

The following version of the Minkowski inequality is also useful.

6.7.53 Lemma (Integral version of Minkowski inequality) ***missing stuff*** *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $I, J \subseteq \mathbb{R}$ *be intervals, and let* $p \in [1, \infty)$*. Let* $f\colon I \times J \to \mathbb{F}$ *have the property that* $x \mapsto f(x, y)$ *is in* $\mathsf{L}^{(p)}(I; \mathbb{F})$ *for almost every* $y \in J$ *and that* $y \mapsto f(x, y)$ *is in* $\mathsf{L}^{(p)}(J; \mathbb{F})$ *for almost every* $x \in I$*. Then, we have*

$$\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^p \mathrm{d}x\Big)^{1/p} \le \int_J \Big(\int_I |f(x,y)|^p \mathrm{d}x\Big)^{1/p} \mathrm{d}y.$$

Proof For $p = 1$ we have

$$\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|\,\mathrm{d}x \le \int_I \Big(\int_J |f(x,y)|\,\mathrm{d}y\Big)\,\mathrm{d}x = \int_J \Big(\int_I |f(x,y)|\,\mathrm{d}x\Big)\,\mathrm{d}y,$$

giving the result in this case by Fubini's Theorem.

Now let $p \in (1, \infty)$. Here we compute

$$\begin{aligned}
\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^p \mathrm{d}x &= \int_I \Big(\Big(\Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p-1}\Big)\Big(\Big|\int_J f(x,z)\,\mathrm{d}z\Big|\Big)\Big)\,\mathrm{d}x \\
&\le \int_I \Big(\int_J \Big(|f(x,z)|\Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p-1}\Big)\,\mathrm{d}z\Big)\,\mathrm{d}x \\
&= \int_J \Big(\int_I \Big(|f(x,z)|\Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p-1}\Big)\,\mathrm{d}x\Big)\,\mathrm{d}z
\end{aligned}$$

using Fubini's Theorem in the last step. Now let $p' = \frac{p}{p-1}$ be the conjugate index. Now, by Hölder's inequality,

$$\begin{aligned}
\int_I \Big(|f(x,z)|\Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p-1}\Big)\,\mathrm{d}x &\le \Big(\int_I |f(x,z)|^p\,\mathrm{d}x\Big)^{1/p}\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p'(p-1)}\mathrm{d}x\Big)^{1/p'} \\
&= \Big(\int_I |f(x,z)|^p\,\mathrm{d}x\Big)^{1/p}\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p}\mathrm{d}x\Big)^{1/p'}.
\end{aligned}$$

Substituting this last relation into the preceding equation yields

$$\begin{aligned}
\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^p \mathrm{d}x &\le \int_J \Big(\Big(\int_I |f(x,z)|^p\,\mathrm{d}x\Big)^{1/p}\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p}\mathrm{d}x\Big)^{1/p'}\Big)\,\mathrm{d}z \\
&= \Big(\int_J \Big(\int_I |f(x,z)|^p\,\mathrm{d}x\Big)^{1/p}\mathrm{d}z\Big)\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p}\mathrm{d}x\Big)^{1/p'}
\end{aligned}$$

Now we note that the lemma is obviously true when

$$\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^p \mathrm{d}x = 0.$$

So we suppose that this quantity is nonzero and divide the above-derived inequality

$$\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^p \mathrm{d}x \le \Big(\int_J \Big(\int_I |f(x,z)|^p\,\mathrm{d}x\Big)^{1/p}\mathrm{d}z\Big)\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^{p}\mathrm{d}x\Big)^{1/p'}$$

by

$$\Big(\int_I \Big|\int_J f(x,y)\,\mathrm{d}y\Big|^p \mathrm{d}x\Big)^{1/p'}$$

which gives the desired inequality after noting that p' is conjugate to p. ■

Now we can prove the basic fact about the spaces $\mathsf{L}^{(p)}(I; \mathbb{F})$.

6.7.54 Proposition (Properties of $(\mathsf{L}^{(p)}(I;\mathbb{F}), \|\cdot\|_\infty)$) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *if* $p \in [1, \infty)$, *and if* $I \subseteq \mathbb{R}$ *is an interval then* $(\mathsf{L}^{(p)}(I;\mathbb{F}), \|\cdot\|_\infty)$ *is a seminormed* $\mathbb{F}$*-vector space. Moreover,* $\|f\|_p = 0$ *if and only if* $f(x) = 0$ *for almost every* $x \in I$.

Proof That $\mathsf{L}^{(p)}(I;\mathbb{F})$ is a seminormed vector space follows from Lemma 6.7.52 which gives the triangle inequality; the other seminorm properties are clear. The final assertion is clear. ■

Now we proceed much as we did for $\mathsf{L}^{(\infty)}(I;\mathbb{F})$. That is, we define

$$Z^p(I;\mathbb{F}) = \{f \in \mathsf{L}^{(p)}(I;\mathbb{F}) \mid \|f\|_p = 0\}$$

and note that, by Theorem 6.1.8, $\mathsf{L}^{(p)}(I;\mathbb{F})/Z^p(I;\mathbb{F})$ is a normed $\mathbb{F}$-vector space if we define the norm by

$$\|f + Z^p(I;\mathbb{F})\|_p = \|f\|_p.$$

This leads to the following definition.

6.7.55 Definition ($\mathsf{L}^p(I;\mathbb{F})$) For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, for $p \in [1, \infty)$, and for an interval $I \subseteq \mathbb{R}$,

$$\mathsf{L}^p(I;\mathbb{F}) = \mathsf{L}^{(p)}(I;\mathbb{F})/Z^p(I;\mathbb{F}).$$ •

We can prove that $\mathsf{L}^p(I;\mathbb{F})$ is a Banach space.

6.7.56 Theorem ($(\mathsf{L}^p(I;\mathbb{F}), \|\cdot\|_p)$ is a Banach space) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *if* $p \in [1, \infty)$, *and if* $I \subseteq \mathbb{R}$ *is an interval, then* $(\mathsf{L}^p(I;\mathbb{F}), \|\cdot\|_p)$ *is an* $\mathbb{F}$*-Banach space. Moreover,* $\mathsf{L}^p(I;\mathbb{F})$ *is isomorphic, as a normed vector space, to the completion of* $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$.

Proof For brevity let us denote $[f] = f + Z^p(I;\mathbb{F})$ for $f \in \mathsf{L}^{(p)}(I;\mathbb{F})$. We use the characterisation of completeness of Theorem 6.4.6. Let $\sum_{j=1}^\infty [f_j]$ be absolutely convergent. Define $g\colon I \to \mathbb{F} \cup \{\infty\}$ by

$$g(x) = \Big(\sum_{j=1}^\infty |f_j(x)|\Big)^p,$$

and note that Minkowski's inequality gives

$$\|g\|_1 \leq \sum_{j=1}^\infty \|f_j\|_p = \sum_{j=1}^\infty \|[f_j]\|_p < \infty$$

since $\sum_{j=1}^\infty [f_j]$ is absolutely convergent. Therefore, $g \in \mathsf{L}^{(1)}(I;\mathbb{F})$, and it, therefore, follows that g is finite for almost every $x \in I$. This implies that for almost every $x \in I$ the series $\sum_{j=1}^\infty f_j(x)$ is absolutely convergent and so convergent. Now define

$$f(x) = \begin{cases} \sum_{j=1}^\infty f_j(x), & g(x) < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Since f is almost everywhere equal to the measurable function g, it is itself measurable, and further $\|f\|_p \leq \|g\|_1 < \infty$ so that $f \in \mathsf{L}^{(p)}(I;\mathbb{F})$. Furthermore, the Dominated

Convergence Theorem gives

$$\begin{aligned}
& f(x) - \sum_{j=1}^{n} f_j(x) = \sum_{j=n+1}^{\infty} f_j(x), \quad \text{a.e. } x \in I \\
\implies \quad & \Big| f(x) - \sum_{j=1}^{n} f_j(x) \Big| \le \sum_{j=n+1}^{\infty} |f_j(x)|, \quad \text{a.e. } x \in I \\
\implies \quad & \lim_{n\to\infty} \Big\| f - \sum_{j=1}^{n} f_j \Big\|_p \le \lim_{n\to\infty} \sum_{j=n+1}^{\infty} \|f_j\|_p = 0 \\
\implies \quad & \lim_{n\to\infty} \Big\| \Big[f - \sum_{j=1}^{n} f_j \Big] \Big\|_p \le \lim_{n\to\infty} \sum_{j=n+1}^{\infty} \|[f_j]\|_p = 0,
\end{aligned}$$

so giving convergence of $\sum_{j=1}^{\infty}[f_j]$.

Now let us prove that $\mathsf{L}^p(I;\mathbb{F})$ is isomorphic, as a normed vector space, to the completion of $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$. We first note that $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ is a subspace of $\mathsf{L}^{(p)}(I;\mathbb{F})$. Moreover, by Exercise 5.9.8 it follows that if $\|f\|_p = 0$ for $f \in \mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ then $f(x) = 0$ for every $x \in I$. That is to say, the map

$$\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F}) \ni f \mapsto [f] \in \mathsf{L}^p(I;\mathbb{F})$$

is injective and so $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ is a subspace of $\mathsf{L}^p(I;\mathbb{F})$. Thus to prove the theorem we need only show that $\mathsf{L}^p(I;\mathbb{F})$ is the closure of $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$. Thus we will show that if $f \in \mathsf{L}^{(p)}(I;\mathbb{F})$ then, for every $\epsilon \in \mathbb{R}_{>0}$ there exists $g \in \mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ such that $\|f - g\|_p < \epsilon$. By Exercise 5.7.4, we can without loss of generality restrict to the case where f takes values in $\mathbb{R}_{\ge 0}$. We shall make this restriction in the arguments below.

Let us first consider the case when $I = [a,b]$ is compact and f is bounded. Let $M \in \mathbb{R}_{>0}$ be such that $f(x) \le M$ for all $x \in I$. Let $\epsilon \in \mathbb{R}_{>0}$. By Theorem 5.9.3 there exists a continuous function $g \colon I \to \mathbb{R}_{\ge 0}$ such that

$$\lambda\Big(\Big\{x \in I \;\Big|\; |f(x) - g(x)| < \tfrac{\epsilon}{(2(b-a))^{1/p}}\Big\}\Big) < \frac{\epsilon^p}{2M^p}.$$

Then

$$\int_a^b |f(x) - g(x)|\, \mathrm{d}x < \frac{\epsilon^p}{2(b-a)}(b-a) + \frac{\epsilon^p}{2M^p} M^p < \epsilon^p.$$

Thus $\|f - g\|_p < \epsilon$, giving the result in this case.

Next we consider the case when $I = [a,b]$ is compact and f is possibly unbounded. Let $\epsilon \in \mathbb{R}_{>0}$. For $M \in \mathbb{R}_{>0}$ define

$$f_M(x) = \begin{cases} f(x), & f(x) \le M, \\ M, & f(x) > M. \end{cases}$$

Since $f \in \mathsf{L}^{(p)}(I;\mathbb{F})$ there exists M sufficiently large that

$$\int_a^b |f(x) - f_M(x)|^p\, \mathrm{d}x < \frac{\epsilon^p}{2^p}.$$

By the argument in the previous paragraph there exists a continuous function $g\colon I \to \mathbb{R}_{\geq 0}$ such that $\|f_M - g\|_p < \frac{\epsilon}{2}$. Then, using the triangle inequality,

$$\|f - g\|_p \leq \|f - f_M\|_p + \|f_M - g\|_p < \epsilon,$$

giving the result in this case.

Finally, we consider the case when I is not compact. Let $\epsilon \in \mathbb{R}_{>0}$. We let $(I_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence of compact intervals such that $I_j \subseteq I_{j+1}$ for each $j \in \mathbb{Z}_{>0}$ and such that $\cup_{j\in\mathbb{Z}_{>0}} I_j = I$. Define a sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{L}^p(I;\mathbb{R})$ by

$$f_j(x) = \begin{cases} f(x), & x \in I_j, \\ 0, & \text{otherwise.} \end{cases}$$

By the Monotone Convergence Theorem we have

$$\lim_{j\to\infty} \int_I |f(x) - f_j(x)|^p \,\mathrm{d}x = \int_I \lim_{j\to\infty} |f(x) - f_j(x)|^p \,\mathrm{d}x = 0.$$

Thus $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to f in $\mathsf{L}^{(p)}(I;\mathbb{F})$. Now, for each $j \in \mathbb{Z}_{>0}$, our arguments above ensure the existence of a continuous function $h_j\colon I_j \to \mathbb{R}_{\geq 0}$ such that $\|f_j|I_j - h_j\|_p^p < \frac{\epsilon^p}{2^{p+1}}$. Note that if we extend h_j to I by asking that it be zero on $I \setminus I_j$ then this extension may not be continuous. However, we can linearly taper h_j to zero on $I \setminus I_j$ to arrive at a continuous function $g_j\colon I \to \mathbb{R}_{\geq 0}$ with compact support satisfying

$$\int_{I\setminus I_j} |g_j(x)|^p \,\mathrm{d}x < \frac{\epsilon^p}{2^{p+1}}.$$

Then

$$\int_I |f_j(x) - g_j(x)|^p \,\mathrm{d}x = \int_{I_j} |f_j(x) - h_j(x)|^p \,\mathrm{d}x + \int_{I\setminus I_j} |g_j(x)|^p \,\mathrm{d}x < \frac{\epsilon^p}{2^{p+1}} + \frac{\epsilon^p}{2^{p+1}} < \frac{\epsilon^p}{2^p}.$$

Now choose $N \in \mathbb{Z}_{>0}$ sufficiently large that $\|f - f_j\|_p < \frac{\epsilon}{2}$. Then, by the triangle inequality,

$$\|f - g_j\|_p \leq \|f - f_j\|_p + \|f_j - g_j\|_p < \epsilon,$$

as desired. ■

6.7.57 Notation (Representing functions in $\mathsf{L}^p(I;\mathbb{F})$) Just as we indicated for $\mathsf{L}^\infty(I;\mathbb{F})$ in Notation 6.7.48, we shall make use of the widespread and convenient convention of identifying an equivalence class in $\mathsf{L}^{(p)}(I;\mathbb{F})$, $p \in [1,\infty)$, with one of its representatives. This is mostly innocuous; however, there are times when this distinction must be made in order for things to make sense. While we do adopt the convention of writing elements of $\mathsf{L}^p(I;\mathbb{F})$ as f rather than $f + Z^p(I;\mathbb{F})$, we shall try to be careful to point out places where it really is the equivalence class that is being used. •

The second part of the Theorem 6.7.56 bears attention. As we commented after the proof of Theorem 6.3.6, although it is not difficult to demonstrate the existence of a completion of a normed vector space, it is not necessarily easy to understand what the meaning of points in the completion are relative to the original normed vector space. The second part of Theorem 6.7.56 says that although elements in the completion of $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ are not functions, they are at least related to functions in that they are equivalence classes of functions. It might also be helpful to view the relationship between $\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ and $\mathsf{L}^p(I;\mathbb{F})$ as being analogous to the relationship between $\mathbb{F}^\infty_0$ and $\ell^p(\mathbb{F})$, as born out in Table 6.2. What is interesting is that, to make this seemingly innocent analogy, one must go through the trials of defining the Lebesgue integral.

Table 6.2 The relationships between the objects in the left column are analogous to the relationships between the objects in the right column

Sequence space	Function space
$\mathbb{F}^\infty_0$	$\mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$
$\ell^p(\mathbb{F})$	$\mathsf{L}^p(I;\mathbb{F})$

Let us prove the separability of $\mathsf{L}^p(I;\mathbb{F})$.

6.7.58 Proposition ($\mathsf{L}^p(I;\mathbb{F})$ is separable) *If* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$, *if* $I \subseteq \mathbb{R}$ *is an interval, and if* $p \in [1,\infty)$, *then* $\mathsf{L}^p(\mathbb{T};\mathbb{F})$ *is separable.*

Proof From Theorem 6.7.40 we know that $\mathsf{C}^0_0(I;\mathbb{F})$ is separable and so $\mathsf{C}_{\mathrm{cpt}}(I;\mathbb{F})$ is also separable, being a subspace of $\mathsf{C}^0_0(I;\mathbb{F})$. Thus a countable dense subset $D \subseteq \mathsf{C}^0_{\mathrm{cpt}}(I;\mathbb{F})$ is also dense in $\mathsf{L}^p(I;\mathbb{F})$ by Exercise 6.6.2. ■

It is useful to be able to relate convergence in $\mathsf{L}^p(I;\mathbb{F})$ to pointwise convergence. The precise statement of this is as follows. Here we are careful to express the result in terms of equivalence classes of functions, since this is important to the meaning of the result. In the statement of the result we denote $[f] = f + Z^p(I;\mathbb{F})$ for brevity.

6.7.59 Proposition (Pointwise convergence and convergence in $\mathsf{L}^p(I;\mathbb{F})$) *Let* $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$, *let* $p \in [1,\infty]$, *and let* $I \subseteq \mathbb{R}$ *be an interval. If* $([f_j])_{j\in\mathbb{Z}_{>0}}$ *is a sequence in* $\mathsf{L}^p(I;\mathbb{F})$ *converging to* $[f] \in \mathsf{L}^p(I;\mathbb{F})$ *then there exists a subsequence* $([f_{j_k}])_{k\in\mathbb{Z}_{>0}}$ *with the property that, for any representatives* $f_{j_k} \in [f_{j_k}]$, $k \in \mathbb{Z}_{>0}$, *and any representative* $f \in [f]_p$, *we have* $\lim_{k\to\infty} f_{j_k}(x) = f(x)$ *for almost every* $x \in I$.

Proof Throughout the proof we work with arbitrary representatives f_{j_k}, $k \in \mathbb{Z}_{>0}$, as stated in the proof. Since $\lim_{j\to\infty}\|f - f_j\|_p = 0$ there exists a subsequence $(f_{j_k})_{k\in\mathbb{Z}_{>0}}$ satisfying $\|f_{j_{k+1}} - f_{j_k}\|_p \le 2^{-k}$. We then define

$$g_k(x) = \sum_{\ell=1}^{k} |f_{j_{k+1}}(x) - f_{j_k}(x)|$$

and $g(x) = \lim_{k\to\infty} g_k(x)$ whenever these quantities are finite, taking them to be zero otherwise. Using Minkowski's inequality, $\|g_k\|_p \le 1$. Fatou's Lemma then gives

$\|g\|_p \leq 1$. This means that $g(x)$ is finite for almost every $x \in I$. Now define

$$f(x) = f_{j_1}(x) + \sum_{j=1}^{\infty} (f_{j_{k+1}}(x) - f_{j_k}(x)) \tag{6.17}$$

when this limit exists, taking it to be zero otherwise. Since the sum converges absolutely for almost every $x \in I$ this implies that the limit in (6.17) exists for almost every $x \in I$. The matter of showing that $f \in \mathsf{L}^p(I; \mathbb{F})$ goes like the last steps in the proof of the completeness of in Theorems 6.7.47 and 6.7.56. This gives the result for a particular representative of the limit class in $\mathsf{L}^p(I; \mathbb{F})$. That the result holds for any representative follows since any two representatives differ on a set of zero measure. ■

6.7.8 Banach spaces of integrable functions on measure spaces

6.7.9 Banach spaces of measures

In this section we let $(X, \mathscr{A})$ be a measurable space, and we recall from Section 5.3.10 the $\mathbb{R}$-vector spaces $\mathsf{M}((X, \mathscr{A}); \mathbb{R})$ and $\mathsf{M}((X, \mathscr{A}); \mathbb{R}^n)$ of finite signed and $\mathbb{R}^n$-valued vector measures on $\mathscr{A}$, and the $\mathbb{C}$-vector space $\mathsf{M}((X, \mathscr{A}); \mathbb{C})$ of complex measures on $\mathscr{A}$. For μ in either $\mathsf{M}((X, \mathscr{A}); \mathbb{R})$ or $\mathsf{M}((X, \mathscr{A}); \mathbb{C})$ the total variation of μ is defined to be

$$\|\mu\| = \sup \Big\{ \sum_{j=1}^{k} |\mu(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } X \Big\}$$

(for signed measures this follows from Proposition 5.3.48). If $\boldsymbol{\mu} \in \mathsf{M}((X, \mathscr{A}); \mathbb{R}^n)$ then the total variation of $\boldsymbol{\mu}$ is defined by

$$|||\boldsymbol{\mu}|||_{\mathbb{R}^n} = \sup \Big\{ \sum_{j=1}^{k} \|\boldsymbol{\mu}(A_j)\|_{\mathbb{R}^n} \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } X \Big\}.$$

We can now state the main result of this section.

6.7.60 Theorem (Banach spaces of measures) *The pairs* $(\mathsf{M}((\mathsf{X}, \mathscr{A}); \mathbb{R}), \|\cdot\|)$ *and* $(\mathsf{M}((\mathsf{X}, \mathscr{A}); \mathbb{R}^{\mathsf{n}}), |||\cdot|||_{\mathbb{R}^{\mathsf{n}}})$ *are* $\mathbb{R}$*-Banach spaces and the pair* $(\mathsf{M}((\mathsf{X}, \mathscr{A}); \mathbb{C}), \|\cdot\|)$ *is a* $\mathbb{C}$*-Banach space.*

Proof We first must verify that $\|\cdot\|$ and $|||\cdot|||_{\mathbb{R}^n}$ are norms. For $\|\cdot\|$, we clearly have $\|\mu\| \in \mathbb{R}_{>0}$ for $\mu \in \mathsf{M}((X, \mathscr{A}); \mathbb{R})$. Also, if $\alpha \in \mathbb{F}$ for $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$,

$$\begin{aligned} \|\alpha\mu\| &= \sup \Big\{ \sum_{j=1}^{k} |\alpha\mu(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } X \Big\} \\ &= |\alpha| \sup \Big\{ \sum_{j=1}^{k} |\mu(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } X \Big\} = |\alpha| \|\mu\|. \end{aligned}$$

If $\mu_1, \mu_2 \in \mathsf{M}((X, \mathscr{A}); \mathbb{F})$ then we have

$$\begin{aligned}\|\mu_1 + \mu_2\| &= \sup\Big\{\sum_{j=1}^{k} |\mu_1(A_j) + \mu_2(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } X\Big\}\\ &\le \sup\Big\{\sum_{j=1}^{k} |\mu_1(A_j)| + |\mu_2(A_j)| \;\Big|\; (A_1, \dots, A_k) \text{ is a partition of } X\Big\}\\ &= \|\mu_1\| + \|\mu_2\|\end{aligned}$$

using Proposition 2.2.27. This gives the triangle inequality for $\|\cdot\|$. Finally, we suppose that $\|\mu\| = 0$. For $A \in \mathscr{A}$ we have

$$|\mu(A)| \le |\mu(A)| + |\mu(X \setminus A)| \le \|\mu\|$$

since $(A, X \setminus A)$ is a partition of X. Thus it follows that $\mu(A) = 0$ for every $A \in \mathscr{A}$. Thus μ is the zero measure. This verifies positive-definiteness of $\|\cdot\|$ and so verifies that it is a norm. An entirely similar analysis yields the same conclusion for $|||\cdot|||_{\mathbb{R}^n}$.

It now remains to verify the completeness of the normed vector spaces. We consider the case of a signed or complex measure, letting $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. We consider a Cauchy sequence $(\mu_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{M}((X, \mathscr{A}); \mathbb{F})$. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that $\|\mu_j - \mu_k\| \le \epsilon$ for $j, k \ge N$. Then, for $A \in \mathscr{A}$, we have, since $(A, X \setminus A)$ is a partition of X,

$$|\mu_j(A) - \mu_k(A)| \le |(\mu_j - \mu_k)(A)| + |(\mu_j - \mu_k)(X \setminus A)| \le \|\mu_j - \mu_k\| \le \epsilon$$

for $j, k \ge N$. Thus, for every $A \in \mathscr{A}$, $(\mu_j(A))_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence in $\mathbb{F}$. We then denote the limit of this Cauchy sequence by $\mu(A)$. We must show that the map $A \mapsto \mu(A)$ is a signed or complex measure.

The following lemma will be useful, saying that the limit $\lim_{j\to\infty} \mu_j(A) = \mu(A)$ in uniform in A.

1 Lemma *For $\epsilon \in \mathbb{R}_{>0}$ there exists $\mathrm{N} \in \mathbb{Z}_{>0}$ such that $|\mu(\mathrm{A}) - \mu_{\mathrm{j}}(\mathrm{A})| < \epsilon$ for each $\mathrm{j} \ge \mathrm{N}$ and $\mathrm{A} \in \mathscr{A}$.*

Proof Let $\epsilon \in \mathbb{R}_{>0}$ and choose $N \in \mathbb{Z}_{>0}$ such that $\|\mu_j - \mu_k\| < \frac{\epsilon}{2}$ for $j, k \ge N$. Thus, as we saw above, $|\mu_j(A) - \mu_k(A)| < \frac{\epsilon}{2}$ for $j, k \ge N$. Now let N_1 be sufficiently large that $|\mu(A) - \mu_k(A)| < \frac{\epsilon}{2}$ for $k \ge N_1$. Now, if $A \in \mathscr{A}$ and $j \ge N$ we have

$$|\mu(A) - \mu_j(A)| \le |\mu(A) - \mu_k(A)| + |\mu_k(A) - \mu_j(A)| < \epsilon,$$

where $k \ge \max\{N, N_1\}$. ▼

Since $\mu_j(\emptyset) = 0$ for every $j \in \mathbb{Z}_{>0}$ we obviously have

$$\mu(\emptyset) = \lim_{j\to\infty} \mu_j(\emptyset) = 0.$$

Let $A_1, \dots, A_m$ be a finite family of pairwise disjoint $\mathscr{A}$-measurable sets. Since μ_j, $j \in \mathbb{Z}_{>0}$, is countably-additive, it is finitely-additive, and so

$$\mu_j(\cup_{l=1}^{m} A_l) = \sum_{l=1}^{m} \mu_j(A_l), \qquad j \in \mathbb{Z}_{>0}.$$

Therefore,

$$\mu(\cup_{l=1}^m A_l) = \lim_{j\to\infty} \mu_j(\cup_{l=1}^m A_l) = \lim_{j\to\infty} \sum_{l=1}^m \mu_j(A_l) = \sum_{l=1}^m \mu(A_l),$$

swapping the finite sum with the limit. This gives finite-additivity of μ. It also holds that μ is consistent since, by construction, it takes values in $\mathbb{R}$.

Now let $(A_l)_{l\in\mathbb{Z}_{>0}}$ be a family of $\mathscr{A}$-measurable sets such that $A_{l+1} \subseteq A_l$, $l \in \mathbb{Z}_{>0}$, and such that $\cap_{l\in\mathbb{Z}_{>0}} A_l = \emptyset$. Since μ_j, $j \in \mathbb{Z}_{>0}$, is countably-additive and consistent, by Proposition 5.3.3 we have

$$\lim_{l\to\infty} \mu_j(A_l) = 0, \qquad j \in \mathbb{Z}_{>0}.$$

Let $\epsilon \in \mathbb{R}_{>0}$ and, by Lemma 1, let $N_1 \in \mathbb{Z}_{>0}$ be such that $|\mu(A) - \mu_j(A)| < \frac{\epsilon}{2}$ for $j \geq N_1$ and $A \in \mathscr{A}$. Let $N \in \mathbb{Z}_{>0}$ be such that $|\mu_{N_1}(A_l)| < \frac{\epsilon}{2}$ for $l \geq N$. Then, for $l \geq N$ we have

$$|\mu(A_l)| \leq |\mu(A_l) - \mu_{N_1}(A_l)| + |\mu_{N_1}(A_l)| < \epsilon.$$

Thus $\lim_{l\to\infty} \mu(A_l) = 0$ and so μ is countable additive by Proposition 5.3.3.

Finally, we must show that $(\mu_j)_{j\in\mathbb{Z}_{>0}}$ converges to μ. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that $\|\mu_j - \mu_k\| < \epsilon$ for $j, k \geq N$. Let $(A_1, \dots, A_m)$ be a partition of X and note that, by definition of $\|\cdot\|$,

$$\sum_{l=1}^m |\mu_j(A_l) - \mu_k(A_l)| = \sum_{l=1}^m |(\mu_j - \mu_k)(A_l)| \leq \|\mu_j - \mu_k\| < \epsilon$$

for $j, k \geq N$. Therefore,

$$\sum_{l=1}^m |\mu(A_l) - \mu_k(A_l)| = \lim_{j\to\infty} \sum_{l=1}^m |\mu_j(A_l) - \mu_k(A_l)| \leq \epsilon$$

for $k \geq N$. Since this holds for every partition $(A_1, \dots, A_m)$ of X, taking the supremum over all such partitions gives $\|\mu - \mu_k\| \leq \epsilon$ for $k \geq N$, so giving convergence of $(\mu_j)_{j\in\mathbb{Z}_{>0}}$ to μ. ■

6.7.10 Notes

Exercises

6.7.1 For $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}_{>0}$ show that

$$\sum_{j=1}^n a_j b_j \leq \max\{b_1, \dots, b_n\} \sum_{j=1}^n a_j.$$

6.7.2 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. For $(a_j)_{j\in\mathbb{Z}_{>0}} \in \ell^1(\mathbb{F})$ and $(b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^\infty(\mathbb{F})$, show that $(a_j b_j)_{j\in\mathbb{Z}_{>0}} \in \ell^1(\mathbb{F})$ and that

$$\|(a_j b_j)_{j\in\mathbb{Z}_{>0}}\|_1 \leq \|(a_j)_{j\in\mathbb{Z}_{>0}}\|_1 \|(b_j)_{j\in\mathbb{Z}_{>0}}\|_\infty.$$

6.7.3 Show that $\mathbb{F}_0^\infty$ is not dense in $\ell^\infty(\mathbb{F})$.

6.7.4 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

(a) Show that $\ell^p(\mathbb{F}) \subseteq \mathsf{c}_0(\mathbb{F})$ for $p \in [1, \infty)$.

(b) Is $\ell^\infty(\mathbb{F}) \subseteq \mathsf{c}_0(\mathbb{F})$?

6.7.5 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $((\mathsf{V}_i, \|\cdot\|_i))_{i \in I}$ be a family of normed F-vector spaces. Show that if $\ell^p(\bigoplus_{i \in I} \mathsf{V}_i)$ is a Banach space for any $p \in [1, \infty)$ then V_i is a Banach space for every $i \in I$.

6.7.6 Show that $\mathsf{C}^0_0(\mathbb{R}; \mathbb{F})$ can be defined alternatively by (6.15).

6.7.7 Show that $\mathsf{C}^0_{\mathrm{cpt}}((0, 1); \mathbb{F})$ is not dense in $\mathsf{L}^\infty((0, 1); \mathbb{R})$.
Hint: *Consider* $\mathrm{f}(\mathrm{x}) = 1$, $\mathrm{x} \in (0, 1)$.

6.7.8 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $I \subseteq \mathbb{R}$ be an interval, and let $f \in \mathsf{L}^{(1)}(I; \mathbb{F})$ and $g \in \mathsf{L}^{(\infty)}(I; \mathbb{F})$. Show that $fg \in \mathsf{L}^{(1)}(I; \mathbb{F})$ and $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$.

Chapter 7

Hilbert spaces

The notion of a Hilbert space is one of the most important in mathematics and applications of mathematics. It will arise in a crucial way in Fourier analysis in Chapters 12, 13, **??**, and 14. Hilbert space theory also plays an important rôle in optimisation theory, system theory, and partial differential equations, to name just as a few applications. As we shall see, Hilbert spaces are examples of Banach spaces, so all of our discussions of Chapter 6 apply to Hilbert spaces. However, the norm in a Hilbert space arises in a particular way, from an inner product. The inner product structure gives rise to important concepts such as orthogonality, and it is concepts such as these that account for the importance of Hilbert spaces as examples of Banach spaces.

In this chapter we give a systematic overview of the notion of a Hilbert space, developing the theory starting in the simple but insightful finite-dimensional case. We endeavour to indicate how all of the concepts in general Banach space theory as developed in Chapter 6 specialise to Hilbert spaces.

Do I need to read this chapter? This chapter is an important one and most of the material in it is essential to the applied material that follows in later volumes. Certain specialised topics can be omitted on an initial reading. In particular, the details of uncountable orthonormal sets in Section 7.3.1 can be initially sidestepped, instead referring explicitly to the countable case considered in Section 7.3.3. •

Contents

7.1 Definitions and properties of inner product spaces 738
7.1.1 Inner products and semi-inner products 738
7.1.2 Inner product spaces as normed vector spaces 741
7.1.3 Orthogonality 746
7.1.4 Hilbert spaces and their subspaces 749
7.1.5 Minimising distance to a set 754
7.1.6 Norms 757
7.2 Continuous maps between inner product spaces 760
7.2.1 The dual of an inner product space 760
7.2.2 Particular aspects of continuity for inner product spaces 763
7.2.3 Notes 764
7.3 Orthonormal bases in Hilbert spaces 765

7.3.1 General definitions and results 765
7.3.2 Finite orthonormal sets and finite Hilbert bases 776
7.3.3 Countable orthonormal sets and countable Hilbert bases 780
7.3.4 Generalised Fourier series 786
7.3.5 Classification of Hilbert spaces 789
7.3.6 Notes 793

Section 7.1

Definitions and properties of inner product spaces

We have already encountered an important example of inner product, the standard inner product on $\mathbb{R}^n$ in Section **??**. The axioms defining a general inner product are exactly those for the standard inner product on $\mathbb{R}^n$, with the slight added generality that we allow for vector spaces over $\mathbb{C}$ as well as over $\mathbb{R}$.

Do I need to read this section? If you are reading this chapter then you should read this section. •

7.1.1 Inner products and semi-inner products

Just as we did in Chapter 6, we will simultaneously deal with the fields $\mathbb{R}$ and $\mathbb{C}$ by letting $\mathbb{F}$ denote wither $\mathbb{R}$ or $\mathbb{C}$, by letting $|a|$, $a \in \mathbb{F}$, denote the absolute value or modulus, and by letting $\bar{a}$, $a \in \mathbb{F}$, denote either a or the complex conjugate of a. We refer to Notation 6.1.1.

Just as in parts of Chapter 6 we considered seminorms, we will also consider semi-inner products in parts of this chapter. There is an additional caveat to make in this respect. The notion of a seminorm has an important independent life separate from its defining a norm as in Theorem 6.1.8. Indeed, in Chapter **??** we will devote significant time and effort to how seminorms arise in linear analysis. However, this is much less the case with the notion of a semi-inner product. Indeed, most authors do not mention the concept. We do so for two reasons: (1) there are examples of semi-inner products that arise *en route* to the construction of certain inner products; (2) we wish to maintain some consistency with the presentation in Chapter 6. Nonetheless, the reader is well-advised to not place much stock in the concept of a semi-inner product and to focus instead on the special case of an inner product.

With all that said, we can give the definitions.

7.1.1 Definition (Semi-inner product, inner product) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be an $\mathbb{F}$-vector space. A ***semi-inner product*** on V is a map $\mathsf{V} \times \mathsf{V} \ni (v_1, v_2) \mapsto \langle v_1, v_2 \rangle \in \mathbb{F}$ with the following properties:

(i) $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle}$ for $v_1, v_2 \in \mathsf{V}$ (***symmetry***);

(ii) $\langle a_1 v_1 + a_2 v_2, v \rangle = a_1 \langle v_1, v \rangle + a_2 \langle v_2, v \rangle$ for $a_1, a_2 \in \mathbb{F}$ and $v_1, v_2 \in \mathsf{V}$ (***linearity***);

(iii) $\langle v, v \rangle \geq 0$ for $v \in \mathsf{V}$, (***positivity***).

An ***inner product*** on V is a semi-inner product $(v_1, v_2) \mapsto \langle v_1, v_2 \rangle$ with the additional property that

(iv) $\langle v, v \rangle = 0$ only if $v = 0_{\mathsf{V}}$ (***definiteness***).

We shall often denote a semi-inner product by $\langle \cdot, \cdot \rangle$. •

Note that the condition for positivity makes sense even when $\mathbb{F} = \mathbb{C}$ since $\langle v, v\rangle$ is always real. Indeed, using symmetry of the semi-inner product,

$$\overline{\langle v, v\rangle} = \overline{\overline{\langle v, v\rangle}} = \langle v, v\rangle,$$

and since the subset $\mathbb{R} \subseteq \mathbb{C}$ is exactly characterised by its being the subset fixed by complex conjugation, it follows that $\langle v, v\rangle \in \mathbb{R}$.

Let us record a trivial consequence of the properties of a semi-inner product.

7.1.2 Proposition (Bilinearity or sesquilinearity of a semi-inner product) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let V be an $\mathbb{F}$-vector space, and let $\langle\cdot,\cdot\rangle$ be a semi-inner product on V. Then, for $\mathrm{a}_1, \mathrm{a}_2, \mathrm{b}_1, \mathrm{b}_2 \in \mathbb{F}$ and $\mathrm{u}_1, \mathrm{u}_2, \mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V}$ we have*

$$\begin{aligned}\langle \mathrm{a}_1\mathrm{u}_1 + \mathrm{a}_2\mathrm{u}_2, \mathrm{b}_1\mathrm{v}_1 + \mathrm{b}_2\mathrm{v}_2\rangle \\ = \mathrm{a}_1\bar{\mathrm{b}}_1\langle \mathrm{u}_1, \mathrm{v}_1\rangle + \mathrm{a}_1\bar{\mathrm{b}}_2\langle \mathrm{u}_1, \mathrm{v}_2\rangle + \mathrm{a}_2\bar{\mathrm{b}}_1\langle \mathrm{u}_2, \mathrm{v}_1\rangle + \mathrm{a}_2\bar{\mathrm{b}}_2\langle \mathrm{u}_2, \mathrm{v}_2\rangle.\end{aligned}$$

Proof We leave this as Exercise 7.1.1. ■

In the case when $\mathbb{F} = \mathbb{R}$ this property is called ***bilinearity*** and when $\mathbb{F} = \mathbb{C}$ this property is called ***sesquilinearity***.

Let us give some examples of inner products and semi-inner products.

7.1.3 Examples (Semi-inner product, inner product)

1. Any $\mathbb{F}$-vector space V has the useless semi-inner product defined by $\langle v_1, v_2\rangle = 0$ for all $v_1, v_2 \in \mathsf{V}$. This is only an inner product in the uninteresting case when $\mathsf{V} = \{0_{\mathsf{V}}\}$.
2. On $\mathbb{F}^n$ define
$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle_2 = \sum_{j=1}^{n} u_j \bar{v}_j.$$
This is readily seen to be an inner product on $\mathbb{F}^n$. In the case when $\mathbb{F} = \mathbb{R}$ this specialises to the standard inner product on $\mathbb{R}^n$ discussed in Section **??**. Note that we use different notation for this object than was used in Chapter **??**, but we will still refer to it as the standard inner product.
3. Recall from Example 4.3.2–**??** that $\mathbb{F}_0^\infty$ denotes the sequences $(a_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{F}$ for which the set $\{j \in \mathbb{Z}_{>0} \mid v_j \neq 0\}$ is finite. Thus sequences in $\mathbb{F}_0^\infty$ are eventually zero. We define
$$\langle (a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}}\rangle_2 = \sum_{j=1}^{\infty} a_j \bar{b}_j,$$
noting that the sum makes sense since it is finite. It is a straightforward exercise to show that $\langle\cdot,\cdot\rangle_2$ is an inner product.
4. Finally, we consider the $\mathbb{F}$-vector space $\mathsf{C}^0([a, b]; \mathbb{F})$ of continuous $\mathbb{F}$-valued functions on the compact interval $[a, b]$. Here we define an inner product on $\mathsf{C}^0([a, b]; \mathbb{F})$ by
$$\langle f, g\rangle = \int_a^b f(x)\bar{g}(x)\,\mathrm{d}x.$$

One readily verifies all properties of the inner product, possibly resorting to Exercise 3.4.1 for the positive-definiteness. ●

Just as all vector spaces were shown to possess a norm in Proposition 6.1.4, we can use a similar strategy to show that all vector spaces possess an inner product.

7.1.4 Proposition (Vector spaces always have at least one inner product) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* V *is an* $\mathbb{F}$*-vector space then there is an inner product on* V*.*

Proof By Theorem 4.3.46 we know the vector space V possesses a basis which establishes an isomorphism ι of V with $\mathbb{F}^J_0$ for some set J. Let us first define an inner product on $\mathbb{F}^J_0$. Writing a typical element of $\mathbb{F}^J_0$ as $(v_j)_{j\in J}$ we define

$$\langle (u_j)_{j\in J}, (v_j)_{j\in J}\rangle_J = \sum_{j\in J} \bar{u}_j v_j,$$

the sum being well-defined since it is finite. To show that $\langle \cdot, \cdot \rangle_J$ is an inner product is a mere matter of checking the definitions. Now define

$$\langle u, v\rangle_{\mathsf{V}} = \langle \iota(u), \iota(v)\rangle_J, \qquad u, v \in \mathsf{V}.$$

To verify that $\langle \cdot, \cdot \rangle_{\mathsf{V}}$ is an inner product is straightforward. Symmetry is obvious. For linearity we compute

$$\langle a_1 v_1 + a_2 v_2, v\rangle_{\mathsf{V}} = \langle \iota(a_1 v_1 + a_2 v_2), v\rangle_J = \iota a_1 \iota(v_1) + a_2 \iota(v_2) v_J = a_1 \langle v_1, v\rangle + a_2 \langle v_2, v\rangle_{\mathsf{V}},$$

using linearity of $\langle \cdot, \cdot \rangle_J$ and ι. Positivity follows immediately from positivity of $\langle \cdot, \cdot \rangle_J$. Definiteness is shown as follows. Suppose that $\langle v, v\rangle_{\mathsf{V}} = 0$. Then $\langle \iota(v), \iota(v)\rangle_J = 0$ and so $\iota(v) = 0_{\mathbb{F}^J_0}$ by definiteness of $\langle \cdot, \cdot \rangle_J$. Thus $v = 0_{\mathsf{V}}$ since ι is an isomorphism. ■

As with the corresponding Proposition 6.1.4, one must take care to understand that the preceding result asserts neither the existence of a unique or even natural inner product. Moreover, there is no assurance that the inner product defined in the preceding result is useful. We refer to Corollary 6.6.27 to see why some vector spaces are incapable of supporting interesting norms; the same idea applies to inner products since, as we shall shortly see, inner products give rise to norms.

Analogous to normed vector spaces we have the following terminology.

7.1.5 Definition (Semi-inner product space, inner product space) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

(i) An $\mathbb{F}$-***semi-inner product space*** is a pair $(\mathsf{V}, \langle \cdot, \cdot \rangle)$ where V is a $\mathbb{F}$-vector space and $\langle \cdot, \cdot \rangle$ is a semi-inner product on V.

(ii) An $\mathbb{F}$-***inner product space*** is a pair $(\mathsf{V}, \langle \cdot, \cdot \rangle)$ where V is a $\mathbb{F}$-vector space and $\langle \cdot, \cdot \rangle$ is an inner product on V. ●

7.1.6 Notation ((Semi-)inner product spaces) As was the case when we were working with seminormed and normed vector spaces, it will be convenient to be able to state results for both semi-inner product spaces and inner product spaces at the same time. In order to facilitate this we will write "(semi-)inner product space" when we wish to mean that either sorts of objects may be used in the statement. ●

7.1.2 Inner product spaces as normed vector spaces

In this section we show that a (semi-)inner product space gives rise in a natural way to an associated (semi)normed vector space. In order to do so, let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a $\mathbb{F}$-(semi-)inner product space and define a map $\mathsf{V} \ni v \mapsto \|v\| \in \mathbb{R}_{\geq 0}$ by $\|v\| = \sqrt{\langle v, v\rangle}$. While we use the notation $\|\cdot\|$ as if this function is a (semi)norm, we do not in fact know that this is a (semi)norm at this point. It is, however, easy to see that $\|\cdot\|$ satisfies all (semi)norm properties except the triangle inequality. In order to verify this we first prove the following result that is of independent interest.

7.1.7 Theorem (Cauchy–Bunyakovsky–Schwarz inequality) *For an* $\mathbb{F}$*-(semi-)inner product space* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *we have*

$$|\langle \mathrm{v}_1, \mathrm{v}_2\rangle| \leq \|\mathrm{v}_1\|\,\|\mathrm{v}_2\|, \qquad \mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V}.$$

Moreover, if $\langle\cdot,\cdot\rangle$ *is an inner product then equality holds in the above expression if and only if* v_1 *and* v_2 *are collinear, i.e., if and only if*

$$\mathrm{span}_{\mathbb{F}}(\mathrm{v}_1) \subseteq \mathrm{span}_{\mathbb{F}}(\mathrm{v}_2) \quad \textit{or} \quad \mathrm{span}_{\mathbb{F}}(\mathrm{v}_2) \subseteq \mathrm{span}_{\mathbb{F}}(\mathrm{v}_1).$$

Proof The result is obviously true for $v_2 = 0$, so we shall suppose that $v_2 \neq 0$. We first prove the result for $\|v_2\| = 1$. In this case we have

$$\begin{aligned}0 &\leq \|v_1 - \langle v_1, v_2\rangle v_2\|^2\\ &= \langle v_1 - \langle v_1, v_2\rangle v_2, v_1 - \langle v_1, v_2\rangle v_2\rangle\\ &= \langle v_1, v_1\rangle - \langle v_1, v_2\rangle\langle v_2, v_1\rangle - \overline{\langle v_1, v_2\rangle}\langle v_1, v_2\rangle + \langle v_1, v_2\rangle\overline{\langle v_1, v_2\rangle}\langle v_2, v_2\rangle\\ &= \|v_1\|^2 - |\langle v_1, v_2\rangle|^2,\end{aligned}$$

where we have used Proposition 7.1.2. Thus we have shown that, provided $\|v_2\| = 1$,

$$|\langle v_1, v_2\rangle|^2 \leq \|v_1\|^2.$$

Taking square roots yields the result in this case. For $\|v_2\| \neq 1$ we define $v_3 = \frac{v_2}{\|v_2\|}$ so that $\|v_3\| = 1$. In this case

$$|\langle v_1, v_3\rangle| \leq \|v_1\| \quad \Longrightarrow \quad \frac{|\langle v_1, v_2\rangle|}{\|v_2\|} \leq \|v_1\|,$$

and so the inequality in the theorem holds.

Note that $\mathrm{span}_{\mathbb{F}}(v_1) \subset \mathrm{span}_{\mathbb{F}}(v_2)$ if and only if $v_1 = 0_{\mathsf{V}}$. In this case it is obvious that equality holds in the stated inequality. Similarly, equality holds if $\mathrm{span}_{\mathbb{F}}(v_2) \subset \mathrm{span}_{\mathbb{F}}(v_1)$. If $\mathrm{span}_{\mathbb{F}}(v_1) = \mathrm{span}_{\mathbb{F}}(v_2)$ then $v_1 = av_2$ for some $a \in \mathbb{F}$. In this case it is a direct computation, using properties of the inner product, to show that the stated inequality is in fact achieved with equality.

Conversely, suppose that the inequality in the theorem is achieved with equality. If equality is achieved with zero on each side then either or both of $\|v_1\|$ and $\|v_2\| = 0$ hold, i.e., either or both of v_1 and v_2 are zero. In this case we have either

$$\mathrm{span}_{\mathbb{F}}(v_1) \subseteq \mathrm{span}_{\mathbb{F}}(v_2) \quad \text{or} \quad \mathrm{span}_{\mathbb{F}}(v_2) \subseteq \mathrm{span}_{\mathbb{F}}(v_1),$$

as desired. Thus the final assertion to prove is that one of the preceding inclusions holds when equality is obtained with both sides of the equality being strictly positive. In this case both of v_1 and v_2 are nonzero. Let us first suppose that $\|v_2\| = 1$. If equality holds in the theorem statement then, going backwards through the argument in the first part of the proof, we must have

$$\|v_1 - \langle v_1, v_2\rangle v_2\|^2 = 0 \qquad \Longrightarrow \qquad v_1 = \langle v_1, v_2\rangle v_2,$$

giving the result in this case. If $\|v_2\| \neq 0$ then define $v_3 = \frac{v_2}{\|v_3\|}$ so that $\|v_3\| = 1$. Moreover,

$$|\langle v_1, v_3\rangle| = \frac{|\langle v_1, v_2\rangle|}{\|v_2\|} = \|v_1\| = \|v_1\|\|v_3\|,$$

and so equality holds for v_1 and v_3 in the inequality in the theorem. By the preceding argument we then have

$$v_1 = \langle v_1, v_3\rangle v_3 \qquad \Longrightarrow \qquad v_1 = \frac{\langle v_1, v_2\rangle}{\|v_2\|^2} v_2,$$

giving the final assertion for $\|v_2\| \neq 0$. ■

Using the Cauchy–Bunyakovsky–Schwarz inequality it is possible to show that the quantity $\|\cdot\|$ associated with an inner product is indeed a norm.

7.1.8 Theorem ((Semi-)inner product spaces are (semi)normed vector spaces) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*-(semi-)inner product space, and define* $\mathsf{V} \ni \mathrm{v} \mapsto \|\mathrm{v}\| \in \mathbb{R}_{\geq 0}$ *be defined by* $\|\mathrm{v}\| = \sqrt{\langle \mathrm{v}, \mathrm{v}\rangle}$. *Then* $(\mathsf{V}, \|\cdot\|)$ *is a (semi)normed vector space.*

Proof All (semi)norm properties except the triangle inequality are easily verified. To verify the triangle inequality, for $v_1, v_2 \in \mathsf{V}$, we compute

$$\begin{aligned}\|v_1 + v_2\|^2 &= \langle v_1 + v_2, v_1 + v_2\rangle = \|v_1\|^2 + \langle v_1, v_2\rangle + \langle v_2, v_1\rangle + \|v_2\|^2 \\ &= \|v_1\|^2 + \langle v_1, v_2\rangle + \overline{\langle v_1, v_2\rangle} + \|v_2\|^2 = \|v_1\|^2 + 2\,\mathrm{Re}(\langle v_1, v_2\rangle) + \|v_2\|^2 \\ &\leq \|v_1\|^2 + 2|\mathrm{Re}(\langle v_1, v_2\rangle)| + \|v_2\|^2 \leq \|v_1\|^2 + 2|\langle v_1, v_2\rangle| + \|v_2\|^2 \\ &\leq \|v_1\|^2 + 2\|v_1\|\|v_2\| + \|v_2\|^2 = (\|v_1\| + \|v_2\|)^2,\end{aligned}$$

using the Cauchy–Bunyakovsky–Schwartz inequality. Taking square roots gives the result. ■

Needless to say, when we talk about the (semi)norm on a (semi-)inner product space, it is the norm of the preceding theorem to which we will refer.

A natural question that arises is then, "Given a norm on a vector space, can one tell when it comes from an inner product?" This question admits an easily stated, but not so easily proved, answer.

7.1.9 Theorem (When does a norm come from an inner product?) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $(\mathsf{V}, \|\cdot\|)$ *is a (semi)normed* $\mathbb{F}$*-vector space, then the following statements are equivalent:*

(i) there exists an (semi-)inner product $\langle\cdot,\cdot\rangle$ *on* V *such that* $\|\mathrm{v}\| = \sqrt{\langle \mathrm{v}, \mathrm{v}\rangle}$ *for all* $\mathrm{v} \in \mathsf{V}$*;*

(ii) $\|\mathrm{v}_1 + \mathrm{v}_2\|^2 + \|\mathrm{v}_1 - \mathrm{v}_2\|^2 = 2\big(\|\mathrm{v}_1\|^2 + \|\mathrm{v}_2\|^2\big)$ *for every* $\mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V}$ *(**parallelogram law**).*

Proof We leave to the reader as Exercise 7.1.4 the fairly easy task of showing that a (semi)norm derived from a (semi-)inner product satisfies the parallelogram law. Here we show the converse.

The proof for $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$ are carried out separately. Let us consider the case of $\mathbb{F} = \mathbb{R}$ first. We claim that if a (semi)norm satisfies the parallelogram law then

$$\langle u, v\rangle \triangleq \tfrac{1}{4}\Big(\|u + v\|^2 - \|u - v\|^2\Big)$$

is a (semi-)inner product on V. It is clear that $\langle u, v\rangle = \langle v, u\rangle$ and that $\langle v, v\rangle \geq 0$ for all $v \in \mathsf{V}$ and that (in the case when $\|\cdot\|$ is a norm) $\langle v, v\rangle = 0$ if and only if $v = 0_{\mathsf{V}}$.

Let $u, v_1, v_2 \in \mathsf{V}$. Then

$$\begin{aligned}\langle u, v_1\rangle + \langle u, v_2\rangle &= \tfrac{1}{4}\Big(\|u + v_1\|^2 - \|u - v_1\|^2 + \|u + v_2\|^2 - \|u - v_2\|^2\Big)\\ &= \tfrac{1}{4}\Big(\Big\|u + \tfrac{1}{2}(v_1 + v_2) + \tfrac{1}{2}(v_1 - v_2)\Big\|^2 - \Big\|u - \tfrac{1}{2}(v_1 + v_2) - \tfrac{1}{2}(v_1 - v_2)\Big\|^2 +\\ &\qquad \Big\|u + \tfrac{1}{2}(v_1 + v_2) - \tfrac{1}{2}(v_1 - v_2)\Big\|^2 - \Big\|u - \tfrac{1}{2}(v_1 + v_2) + \tfrac{1}{2}(v_1 - v_2)\Big\|^2\Big).\end{aligned} \tag{7.1}$$

By the parallelogram law we have

$$\begin{aligned}\Big\|u + \tfrac{1}{2}(v_1 + v_2) + \tfrac{1}{2}(v_1 - v_2)\Big\|^2 + \Big\|u + \tfrac{1}{2}(v_1 + v_2) - \tfrac{1}{2}(v_1 - v_2)\Big\|^2 =&\\ 2\Big\|u + \tfrac{1}{2}(v_1 + v_2)\Big\|^2 + 2\Big\|\tfrac{1}{2}(v_1 - v_2)\Big\|^2&\end{aligned} \tag{7.2}$$

and

$$\begin{aligned}\Big\|u - \tfrac{1}{2}(v_1 + v_2) + \tfrac{1}{2}(v_1 - v_2)\Big\|^2 + \Big\|u - \tfrac{1}{2}(v_1 + v_2) - \tfrac{1}{2}(v_1 - v_2)\Big\|^2 =&\\ 2\Big\|u - \tfrac{1}{2}(v_1 + v_2)\Big\|^2 + 2\Big\|\tfrac{1}{2}(v_1 - v_2)\Big\|^2.&\end{aligned} \tag{7.3}$$

If we substitute (7.2) and (7.3) into (7.1) we get

$$\langle u, v_1\rangle + \langle u, v_2\rangle = \tfrac{1}{4}\Big(\Big\|u + \tfrac{1}{2}(v_1 + v_2)\Big\|^2 - \Big\|u - \tfrac{1}{2}(v_1 + v_2)\Big\|^2\Big) = 2\Big\langle u, \tfrac{1}{2}(v_1 + v_2)\Big\rangle. \tag{7.4}$$

With this we prove a lemma.

1 **Lemma** *If* $\mathrm{k} \in \mathbb{Z}_{\geq 0}$ *then* $\Big\langle \frac{1}{2^{\mathrm{k}}}\mathrm{u}, \mathrm{v}\Big\rangle = \frac{1}{2^{\mathrm{k}}}\langle \mathrm{u}, \mathrm{v}\rangle$ *for all* $\mathrm{u}, \mathrm{v} \in \mathsf{V}$.

Proof The result is vacuously true for $k = 0$. If we let $v_2 = 0$ in (7.4) we have $\Big\langle \frac{1}{2}u, v\Big\rangle = \frac{1}{2}\langle u, v\rangle$, giving the lemma for $k = 1$. Now we proceed by induction. Suppose that the lemma holds for $k = m \geq 2$. Then

$$\Big\langle \tfrac{1}{2^{m+1}}u, v\Big\rangle = \Big\langle \tfrac{1}{2^m}\tfrac{1}{2}u, v\Big\rangle = \tfrac{1}{2^m}\Big\langle \tfrac{1}{2}u, v\Big\rangle = \tfrac{1}{2^{m+1}}\langle u, v\rangle,$$

using the induction hypotheses. ▼

Note that we now have

$$\langle u, v_1 + v_2\rangle = \langle v_1 + v_2, u\rangle = 2\Big\langle \tfrac{1}{2}(v_1 + v_1), u\Big\rangle = 2\Big\langle u, \tfrac{1}{2}(v_1 + v_2)\Big\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle$$

where we have used (7.4).

Now we give another lemma.

2 Lemma *We have* $\left\langle \frac{\mathrm{m}}{2^{\mathrm{k}}}\mathrm{u}, \mathrm{v}\right\rangle = \frac{\mathrm{m}}{2^{\mathrm{k}}}\langle \mathrm{u}, \mathrm{v}\rangle$ *for all* $\mathrm{u}, \mathrm{v} \in \mathsf{V}$, $\mathrm{m} \in \mathbb{Z}$*, and* $\mathrm{k} \in \mathbb{Z}_{\geq 0}$.

Proof We shall prove the result for $m \in \mathbb{Z}_{>0}$. The result for $m = 0$ is trivial, and the proof for $m \in \mathbb{Z}_{<0}$ follows along the same lines as the proof for $m \in \mathbb{Z}_{>0}$.

The result is clearly true for $m = 1$. Now suppose it is true for $m = l \geq 2$. Then we have

$$\begin{aligned} \left\langle \tfrac{l+1}{2^k}u, v\right\rangle = \left\langle \tfrac{l+1}{2^k}u, v\right\rangle &= \left\langle \tfrac{l}{2^k}u + \tfrac{1}{2^k}u, v\right\rangle \\ &= \left\langle \tfrac{l}{2^k}u, v\right\rangle + \left\langle \tfrac{1}{2^k}u, v\right\rangle = \left\langle \tfrac{l}{2^k}u, v\right\rangle + \left\langle \tfrac{1}{2^k}u, v\right\rangle \\ &= \tfrac{l}{2^k}\langle u, v\rangle + \tfrac{1}{2^k}\langle u, v\rangle = \tfrac{l+1}{2^k}\langle u, v\rangle, \end{aligned}$$

using the induction hypotheses. ▼

Now need a pair of technical lemmata.

3 Lemma *Let* $\mathrm{a}, \mathrm{b} \in \mathbb{R}$ *be such that* $\mathrm{a} < \mathrm{b}$. *Then there exist* $\mathrm{m} \in \mathbb{Z}$ *and* $\mathrm{k} \in \mathbb{Z}_{\geq 0}$ *such that* $\mathrm{a} < \frac{\mathrm{m}}{2^{\mathrm{k}}} < \mathrm{b}$.

Proof This is Exercise 2.1.5. ▼

4 Lemma *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \|\cdot\|)$ *be a (semi)normed* $\mathbb{F}$*-vector space. Fix* $\mathrm{u}, \mathrm{v} \in \mathsf{V}$ *and define* $\phi\colon \mathbb{F} \to \mathbb{R}$ *by* $\phi(\mathrm{a}) = \|\mathrm{au} + \mathrm{v}\|$. *Then* ϕ *is continuous.*

Proof This follows from Proposition 6.5.4 along with the fact that the composition of continuous maps is continuous. ▼

We may now prove the final property needed to show that $\langle \cdot, \cdot \rangle$ is a (semi-)inner product. That is, we show that $\langle au, v\rangle = a\langle u, v\rangle$ for all $a \in \mathbb{F}$ and $u, v \in \mathsf{V}$. We will show that $|\langle au, v\rangle - a\langle u, v\rangle| < \epsilon$ for any $\epsilon \in \mathbb{R}_{>0}$. Let $\delta_{m,k} = a - m/2^k$ for $m \in \mathbb{Z}$ and $k \in \mathbb{Z}_{\geq 0}$. Note that we can make $|\delta_{m,k}|$ as small as we like by appropriately choosing m and k. We thus have

$$\begin{aligned} |\langle au, v\rangle - a\langle u, v\rangle| &= |\langle (m2^n + \delta_{m,k})u, v\rangle - (m/2^n + \delta_{m,k})\langle u, v\rangle| \\ &= |\langle \delta_{m,k}u, v\rangle - \delta_{m,k}\langle u, v\rangle| \leq |\langle \delta_{m,k}u, v\rangle| + |\delta_{m,k}\langle u, v\rangle|. \end{aligned}$$

For $\epsilon > 0$ let $\delta_1 = \left|\frac{\epsilon}{2\langle u, v\rangle}\right|$ and let δ_2 be such that $|\langle \delta_2 u, v\rangle| \leq \epsilon/2$. This is possible since $a \mapsto \langle au, v\rangle$ is continuous by Proposition 7.2.1. Now choose $m \in \mathbb{Z}$ and $k \in \mathbb{Z}_{>0}$ so that $\delta_{m,k} < \min(\delta_1, \delta_2)$. Then

$$|\langle au, v\rangle - a\langle u, v\rangle| \leq |\langle \delta_{m,k}u, v\rangle| + |\delta_{m,k}\langle u, v\rangle| < \epsilon/2 + \epsilon/2 = \epsilon,$$

as desired. This shows that $\langle \cdot, \cdot \rangle$ is a (semi-)inner product. Now we show that $\|\cdot\|$ is derived from this (semi-)inner product. This is easy since

$$\langle v, v\rangle = \tfrac{1}{4}\|v + v\|^2 = \|v\|^2.$$

This completes the proof for the case when $\mathbb{F} = \mathbb{R}$.

When $\mathbb{F} = \mathbb{C}$ we claim that

$$\langle u, v\rangle \triangleq \tfrac{1}{4}\left(\|u + v\|^2 - \|u - v\|^2\right) + \tfrac{\mathrm{i}}{4}\left(\|u + \mathrm{i}v\|^2 - \|u - \mathrm{i}v\|^2\right)$$

is a (semi-)inner product on V. First note that

$$\begin{aligned}\overline{\langle v,u\rangle} &= \tfrac{1}{4}\Big(\|v+u\|^2 - \|v-u\|^2\Big) - \tfrac{\mathrm{i}}{4}\Big(\|v+\mathrm{i}u\|^2 - \|v-\mathrm{i}u\|^2\Big)\\ &= \tfrac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2\Big) - \tfrac{\mathrm{i}}{4}\Big(\|{-\mathrm{i}\mathrm{i}v}+\mathrm{i}u\|^2 - \|{-\mathrm{i}\mathrm{i}v}-\mathrm{i}u\|^2\Big)\\ &= \tfrac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2\Big) + \tfrac{\mathrm{i}}{4}\Big(\|\mathrm{i}\mathrm{i}v+\mathrm{i}u\|^2 - \|\mathrm{i}\mathrm{i}v-\mathrm{i}u\|^2\Big)\\ &= \tfrac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2\Big) + \tfrac{\mathrm{i}}{4}\Big(\|u+\mathrm{i}v\|^2 - \|u-\mathrm{i}v\|^2\Big)\\ &= \langle u,v\rangle.\end{aligned}$$

We also compute

$$\langle u, v_1+v_2\rangle = \langle u,v_1\rangle + \langle u,v_2\rangle, \quad \langle au,v\rangle = a\langle u,v\rangle$$

for $u,v,v_1,v_2 \in \mathsf{V}$ and for $a \in \mathbb{R}$. We also compute

$$\begin{aligned}\langle \mathrm{i}u,v\rangle &= \tfrac{1}{4}\Big(\|\mathrm{i}u+v\|^2 - \|\mathrm{i}u-v\|^2\Big) + \tfrac{\mathrm{i}}{4}\Big(\|\mathrm{i}u+\mathrm{i}v\|^2 - \|\mathrm{i}u-\mathrm{i}v\|^2\Big)\\ &= \tfrac{\mathrm{i}}{4}\Big(\|u+v\|^2 - \|u-v\|^2\Big) + \tfrac{1}{4}\Big(\|\mathrm{i}u-\mathrm{i}\mathrm{i}v\| - \|\mathrm{i}u+\mathrm{i}\mathrm{i}v\|\Big)\\ &= \mathrm{i}\tfrac{1}{4}\Big(\|u+v\|^2 - \|u-v\|^2\Big) + \mathrm{i}\tfrac{\mathrm{i}}{4}\Big(\|u+\mathrm{i}v\|^2 - \|u-\mathrm{i}v\|^2\Big)\\ &= \mathrm{i}\langle u,v\rangle\end{aligned}$$

We can then readily check that $\langle au,v\rangle = a\langle u,v\rangle$ for every $u,v \in \mathsf{V}$ and $a \in \mathbb{C}$. This shows that $\langle\cdot,\cdot\rangle$ is a (semi-)inner product. We also have

$$\langle v,v\rangle = \tfrac{1}{4}\|2v\|^2 + \tfrac{\mathrm{i}}{4}|1+\mathrm{i}|^2\|v\|^2 - |1-\mathrm{i}|^2\|v\|^2 = \tfrac{1}{4}\|2v\|^2 = \|v\|^2.$$

Taking square roots shows that $\|\cdot\|$ is the (semi)norm derived from the inner product $\langle\cdot,\cdot\rangle$, and so gives the theorem when $\mathbb{F} = \mathbb{C}$. ■

As a consequence of the proof we have the following formulae which relate an inner product to the norm defined by it.

7.1.10 Corollary (Polarisation identity) *Let $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ and let $(\mathsf{V},\langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-(semi-)inner product space with $\|\cdot\|$ the norm defined by $\langle\cdot,\cdot\rangle$. The following statements hold:*

(i) if $\mathbb{F} = \mathbb{R}$ then

$$\langle \mathrm{u},\mathrm{v}\rangle = \tfrac{1}{4}\Big(\|\mathrm{u}+\mathrm{v}\|^2 - \|\mathrm{u}-\mathrm{v}\|^2\Big)$$

for all $\mathrm{u},\mathrm{v} \in \mathsf{V}$;

(ii) if $\mathbb{F} = \mathbb{C}$ then

$$\langle \mathrm{u},\mathrm{v}\rangle \triangleq \tfrac{1}{4}\Big(\|\mathrm{u}+\mathrm{v}\|^2 - \|\mathrm{u}-\mathrm{v}\|^2\Big) + \tfrac{\mathrm{i}}{4}\Big(\|\mathrm{u}+\mathrm{i}\mathrm{v}\|^2 - \|\mathrm{u}-\mathrm{i}\mathrm{v}\|^2\Big)$$

for all $\mathrm{u},\mathrm{v} \in \mathsf{V}$.

The fact is that it is unusual for a norm to be derived from an inner product. However, since norms coming from inner products are so important, we will devote a great deal of effort to this special case.

With (semi-)inner product spaces now being normed vector spaces, all the norm machinery can be piled into a (semi-)inner product space. Indeed, we shall in this chapter freely refer to any part of Chapter 6. Also, we shall frequently apply the name for a (semi)normed vector space concepts directly to a (semi-)inner product space.

7.1.3 Orthogonality

One of the essential features of an inner product spaces that distinguish them from more general normed vector spaces is that one has the notion of orthogonality. We have some intuition about what orthogonality means in low dimensions (see Section **??**), and some of this intuition carries over to general inner product spaces. However, as is often the case when one makes the leap to infinite-dimensions, one must be careful in relying solely on intuition in making assertions about what is true or not.

Let us give the definitions. Note that the word "orthogonal" has multiple meanings, depending on context.

7.1.11 Definition (Orthogonal, orthogonal complement) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-(semi-)inner product space.

(i) Vectors $v_1, v_2 \in \mathsf{V}$ are ***orthogonal*** if $\langle v_1, v_2\rangle = 0$. We shall write $v_1 \perp v_2$ to denote v_1 and v_2 being orthogonal.

(ii) Sets $A_1, A_2 \subseteq \mathsf{V}$ are ***orthogonal*** if $\langle v_1, v_2\rangle = 0$ for every $v_1 \in A_1$ and $v_2 \in A_2$. We shall write $A_1 \perp A_2$ to denote A_1 and A_2 being orthogonal.

(iii) If $A \subseteq \mathsf{V}$ then the ***orthogonal complement*** of A is the set

$$A^{\perp} = \{u \in \mathsf{V} \mid \langle u, v\rangle = 0 \text{ for all } v \in A\}. \qquad \bullet$$

Let us give some elementary examples of orthogonal sets.

7.1.12 Examples (Orthogonality)

1. The vectors $(1, 2\mathrm{i}, -1), (1, \frac{3}{2} + \frac{\mathrm{i}}{2}, 3\mathrm{i}) \in \mathbb{C}^3$ are orthogonal.
2. In $\mathbb{F}^3$ the sets

$$A_1 = \operatorname{span}_{\mathbb{F}}((1,1,1),(0,1,1)), \quad A_2 = \operatorname{span}_{\mathbb{F}}((0,1,-1))$$

are orthogonal. Moreover, A_1 is the orthogonal complement of A_2 and A_2 is the orthogonal complement of A_1. •

In some sense, this entire chapter is about orthogonality. Let us here give a few simple consequences of the definitions.

7.1.13 Proposition (Properties of orthogonal complement) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*(semi-)-inner product space, and let* $\mathrm{A}, \mathrm{B} \subseteq \mathsf{V}$*. Then the following statements hold:*

(i) if $\mathrm{A} \subseteq \mathrm{B}$ *then* $\mathrm{B}^{\perp} \subseteq \mathrm{A}^{\perp}$*;*

(ii) $\mathrm{A} \subseteq (\mathrm{A}^{\perp})^{\perp}$*;*

(iii) $\mathrm{A}^{\perp}$ *is a closed subspace of* V*;*

(iv) $\mathrm{A}^{\perp} = (\mathrm{cl}(\operatorname{span}_{\mathbb{F}}(\mathrm{A})))^{\perp}$*.*

If $\langle\cdot,\cdot\rangle$ *is additionally an inner product then*

(v) $\mathrm{cl}(\operatorname{span}_{\mathbb{F}}(\mathrm{A})) \cap \mathrm{A}^{\perp} = \{0_{\mathsf{V}}\}$*.*

Proof The proof of parts (i), (ii), and (iii) are left to the reader as Exercise 7.1.11.

(iv) Since $A \subseteq \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A))$ it follows from part (i) that

$$A^{\perp} \supseteq (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^{\perp}.$$

Now let $u \in A^{\perp}$ so that $\langle u, v\rangle = 0$ for every $v \in A$. Next let $\hat{v} \in \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A))$ and by Proposition 6.6.8 let $(\hat{v}_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathrm{span}_{\mathbb{F}}(A)$ converging to $\hat{v}$. For each $j \in \mathbb{Z}_{>0}$ we can write

$$\hat{v}_j = \sum_{r=1}^{k_j} c_{jr} v_{jr}$$

for some $k_j \in \mathbb{Z}_{>0}$ and $c_{jr} \in \mathbb{F}$ and $v_{jr} \in A$, $r \in \{1, \dots, k_j\}$. It therefore follows that

$$\left\langle u, \hat{v}_j \right\rangle = \left\langle u, \sum_{r=1}^{k_j} c_{jr} v_{jr} \right\rangle = \sum_{j=1}^{k_j} \bar{c}_{jr} \langle u, v_{jr}\rangle = 0$$

for each $j \in \mathbb{Z}_{>0}$. This allows us to deduce that

$$\left\langle u, \hat{v} \right\rangle = \lim_{j\to\infty} \left\langle u, \hat{v}_j \right\rangle = 0$$

by Proposition 7.2.1 and Theorem 6.5.2. Thus $u \in (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^{\perp}$ as desired.

(v) If

$$v \in \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap A^{\perp} = \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^{\perp}$$

then $\langle v, v\rangle = 0$ which gives $v = 0_{\mathsf{V}}$ if $\langle \cdot, \cdot \rangle$ is an inner product. ■

The equality $A^{\perp} = (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^{\perp}$ is an important one. It tells us that the orthogonal complement of a set is not a feature of the set, but of the closure of the subspace generated by this set. Thus there are two operations happening when taking orthogonal complements: "span" and "closure" (in that order). The appearance of the topological closure operation here is perhaps surprising at first encounter. Indeed, since all subspaces are closed in finite dimensions, closure does not make an appearance in that case.

It is fairly obviously true that $A \neq (A^{\perp})^{\perp}$ in general, merely because A may not be a subspace but $(A^{\perp})^{\perp}$ is a subspace. So the question of when $A = (A^{\perp})^{\perp}$ is only interesting when A is a subspace. However, even in this case equality does not generally hold. This is something that we will explore in greater detail in ***missing stuff***, so here we merely content ourselves with a counterexample.

7.1.14 Example ($\mathsf{U} \neq (\mathsf{U}^{\perp})^{\perp}$) Let us take $\mathsf{V} = \ell^2(\mathbb{F})$ with the inner product

$$\langle (a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}} \rangle = \sum_{j=1}^{\infty} a_j \bar{b}_j.$$

This is a specialisation to $p = 2$ of the Banach space $\ell^p(\mathbb{F})$ considered in Section 6.7.2. We showed in Theorem 6.7.19 that this is a Banach space and in Corollary 6.7.21 that this Banach space is the completion of $\mathbb{F}_0^{\infty}$. Let us then take the subspace $\mathbb{F}_0^{\infty}$

of $\ell^2(\mathbb{F})$. We claim that $\mathbb{F}_0^\infty$ is a strict subspace of $((\ell_0^\infty)^\perp)^\perp$. To see this we first claim that $(\mathbb{F}_0^\infty)^\perp = \{0_{\ell^2(\mathbb{F})}\}$. Indeed, let $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ be the standard basis for $\mathbb{F}_0^\infty$. Thus, as a reminder,

$$\boldsymbol{e}_j(k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

Then, if $(a_j)_{j\in\mathbb{Z}_{>0}} \in (\mathbb{F}_0^\infty)^\perp$ then

$$\langle (a_j)_{j\in\mathbb{Z}_{>0}}, \boldsymbol{e}_k\rangle = a_k = 0$$

for every $k \in \mathbb{Z}_{>0}$. Thus $(\mathbb{F}_0^\infty)^\perp = \{0_{\ell^2(\mathbb{F})}\}$ as claimed. It, therefore, follows that $((\mathbb{F}_0^\infty)^\perp)^\perp = \ell^2(\mathbb{F})$ and so we have $\mathbb{F}_0^\infty$ as a strict subspace of $((\mathbb{F}_0^\infty)^\perp)^\perp$ as claimed. •

The issue with the preceding example, as we shall see in Theorem 7.1.19, is that $\mathbb{F}_0^\infty$ is not a *closed* subspace of $\ell^2(\mathbb{F})$.

Let us also record how orthogonality interacts with sums and intersections of subsets of V. For $A, B \subseteq$ V we denote

$$A + B \triangleq \{u + v \mid u \in A,\ v \in B\}.$$

We now have the following assertions.

7.1.15 Proposition (Orthogonality and sum and intersection) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be a $\mathbb{F}$-inner product space, and let* $\mathrm{A}, \mathrm{B} \subseteq \mathsf{V}$*. Then the following statements hold:*

(i) $(\mathrm{A} + \mathrm{B})^\perp = \mathrm{A}^\perp \cap \mathrm{B}^\perp$.

(ii) $(\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(\mathrm{A})) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(\mathrm{B})))^\perp = \mathrm{A}^\perp + \mathrm{B}^\perp$.

Proof (i) By part (iv) of Proposition 7.1.13 we have

$$(A + B)^\perp = (\mathrm{span}_{\mathbb{F}}(A + B))^\perp = (\mathrm{span}_{\mathbb{F}}(A) + \mathrm{span}_{\mathbb{F}}(B))^\perp,$$

using the easily verified identity $\mathrm{span}_{\mathbb{F}}(A+B) = \mathrm{span}_{\mathbb{F}}(A) + \mathrm{span}_{\mathbb{F}}(B)$. Let $w \in (A+B)^\perp$. Then $\langle w, u + v\rangle = 0$ for every $u \in \mathrm{span}_{\mathbb{F}}(A)$ and $v \in \mathrm{span}_{\mathbb{F}}(B)$. In particular, $\langle w, u\rangle = 0$ and $\langle w, v\rangle = 0$ for every $u \in \mathrm{span}_{\mathbb{F}}(A)$ and $v \in \mathrm{span}_{\mathbb{F}}(B)$. Thus $w \in A^\perp \cap B^\perp$. Next suppose that $w \in A^\perp \cap B^\perp$. Then, using part (iv) of Proposition 7.1.13,

$$w \in (\mathrm{span}_{\mathbb{F}}(A))^\perp \cap (\mathrm{span}_{\mathbb{F}}(B))^\perp.$$

Therefore, $\langle w, u\rangle = \langle w, v\rangle = 0$ for every $u \in \mathrm{span}_{\mathbb{F}}(A)$ and $v \in \mathrm{span}_{\mathbb{F}}(B)$. Thus $\langle w, u+v\rangle = 0$ $u \in \mathrm{span}_{\mathbb{F}}(A)$ and $v \in \mathrm{span}_{\mathbb{F}}(B)$, giving

$$w \in (\mathrm{span}_{\mathbb{F}}(A) + \mathrm{span}_{\mathbb{F}}(B))^\perp = (A + B)^\perp,$$

as desired.

(ii)

Conversely, since

$$\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \subseteq \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A))$$

we have

$$(\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^\perp \subseteq (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^\perp$$

by part (i) of Proposition 7.1.13. By part (iv) of the same result we then have

$$A^\perp \subseteq (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^\perp.$$

In like manner

$$B^\perp \subseteq (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^\perp.$$

Since $(\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^\perp$ is a subspace by part (iii) of Proposition 7.1.13 it then follows that

$$A^\perp + B^\perp \subseteq (\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)) \cap \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(A)))^\perp,$$

giving the desired conclusion. ■

7.1.4 Hilbert spaces and their subspaces

As inner-product spaces are normed vector spaces, the whole discussion of Cauchy sequences, convergent sequences, and completeness in Sections 6.2 and 6.3 can be applied to inner product spaces. The notion of a complete inner product space is important enough to have its own name.

7.1.16 Definition (Completeness, Hilbert space) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. A $\mathbb{F}$-inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is ***complete*** if the corresponding normed vector space is complete. A ***$\mathbb{F}$-Hilbert***[1] ***space*** is a complete $\mathbb{F}$-inner product space. •

Since inner product spaces are also normed vector spaces, the construction of the completion in Theorem 6.3.6 also applies to inner product spaces. That is to say, every inner product space possesses a completion that is a Banach space. Of course, one would also like to have the completion be a Hilbert space, and this is the content of the next result.

7.1.17 Theorem (Completion of an inner product space) *If* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is an inner product space then there exists a Hilbert space* $(\overline{\mathsf{V}}, \overline{\langle\cdot,\cdot\rangle})$ *and an injective linear map* $\mathrm{i}_{\mathsf{V}}\colon \mathsf{V} \to \overline{\mathsf{V}}$ *with the following properties:*

(i) $\mathrm{image}(\mathrm{i}_{\mathsf{V}})$ *is dense in* $\overline{\mathsf{V}}$*;*

(ii) $\langle \mathrm{v}_1, \mathrm{v}_2\rangle = \overline{\langle \mathrm{i}_{\mathsf{V}}(\mathrm{v}_1), \mathrm{i}_{\mathsf{V}}(\mathrm{v}_2)\rangle}$.

Proof We let $\overline{\mathsf{V}}$ be the vector space constructed from the normed vector space associated to V as in Theorem 6.3.6, and we let i_{V} also be the linear map constructed in the proof of that result. Now let $\overline{v} \in \overline{\mathsf{V}}$ and $v \in \mathsf{V}$ and let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in V for which $\overline{v} = \lim_{j\to\infty} v_j$. We claim that the sequence $(\langle v_j, v\rangle)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{F}$ converges. We may suppose that $v \neq 0$ without loss of generality. Let $\epsilon > 0$ and choose $N \in \mathbb{Z}_{>0}$ so that $\|v_j - v_k\| < \frac{\epsilon}{\|v\|}$ for $j, k \geq N$; this is possible by continuity of the norm. By the Cauchy–Bunyakovsky–Schwarz inequality we then have

$$|\langle v_j, v\rangle - \langle v_k, v\rangle| \leq |\langle v_j - v_k, v\rangle| \leq \|v_j - v_k\|\|v\| \leq \epsilon$$

[1] David Hilbert (1862–1943) in one of history's greatest mathematicians. At the 1900 International Congress of Mathematics in Paris, Hilbert gave a list of twenty three problems which he felt should guide mathematical research in the upcoming centuries. Many of Hilbert's problems have been solved, some to great aplomb. Hilbert's own contributions were in many fields, including geometry, analysis, logic, and algebra.

for $j,k \geq N$, showing that $(\langle v_j, v\rangle)_{j\in\mathbb{Z}_{>0}}$ is Cauchy, and so convergent. Thus we may sensibly define $\overline{\langle \overline{v}, v\rangle} = \lim_{j\to\infty}\langle v_j, v\rangle$. We may similarly, of course, define $\overline{\langle v, \overline{v}\rangle}$, thus defining $\overline{\langle\cdot,\cdot\rangle}$ on $\overline{\mathsf{V}}\times\mathsf{V}$ and $\mathsf{V}\times\overline{\mathsf{V}}$. The same sort of arguments also allow one to define $\overline{\langle\overline{v}_1,\overline{v}_2\rangle}$ for $\overline{v}_1,\overline{v}_2 \in \overline{\mathsf{V}}$. To show that the resulting map $\overline{\mathsf{V}}\times\overline{\mathsf{V}} \ni (\overline{v}_1,\overline{v}_2) \mapsto \overline{\langle\overline{v}_1,\overline{v}_2\rangle} \in \mathbb{F}$ is an inner product is a simple verification of the axioms, using the fact, for example, that if a sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ converges to v, then the sequence $(av_j)_{j\in\mathbb{Z}_{>0}}$ converges to av for $a \in \mathbb{F}$. That (i) holds is an immediate consequence of Theorem 6.3.6, and (ii) is obvious. ■

Let us consider our inner product space examples to determine which are Hilbert spaces.

7.1.18 Examples (Hilbert spaces and non-Hilbert spaces)

1. The inner product space $(\mathbb{F}^n, \langle\cdot,\cdot\rangle_2)$ is a Banach space by virtue of the fact that every finite-dimensional inner product space is complete (Theorem 6.3.3).
2. The inner product space $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$ is not complete. Indeed, in Corollary 6.7.21 we saw that its completion is $\ell^2(\mathbb{F})$ which contains $\mathbb{F}_0^\infty$ as a strict subset. To "by hand" show that $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$ is not complete can be done following the strategy of Example 6.3.1–1. We leave the working out of this to the reader as Exercise 7.1.6. The completion of $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$ is $(\ell^2(\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ as is proved in Corollary 6.7.21.
3. The inner product space $(\mathsf{C}^0([a,b];\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ is not a Hilbert space if $b > a$. In ***missing stuff*** we showed that $\mathsf{L}^2([a,b];\mathbb{F})$ is the completion of $\mathsf{C}^0([a,b];\mathbb{F})$ with respect to the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Since $\mathsf{C}^0([a,b];\mathbb{F})$ is a strict subset of $\mathsf{L}^2([a,b];\mathbb{F})$ this allows us to conclude that $(\mathsf{C}^0([a,b];\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ is not complete. Moreover, one can show this explicitly following the arguments of Example 6.3.1–2; see Exercise 7.1.7. •

The following conclusion for complete subspaces of inner product spaces is important. Note that definiteness of the inner product is essential here.

7.1.19 Theorem (Complete subspaces and direct sum decompositions) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is an $\mathbb{F}$-inner product space, and if U is a complete subspace of V, then $\mathsf{V} = \mathsf{U}\oplus\mathsf{U}^\perp$.*

Proof We refer ahead to Theorem 7.1.25 for a characterisation of the minimisation of the distance from a point to a convex subset. There is nothing in that theorem that involves machinery not yet available to us.

Let $v_0 \in \mathsf{V}$ and, by Theorem 7.1.25, let $\hat{v}_0 \in \mathsf{U}$ be the unique vector such that

$$\|v_0 - \hat{v}_0\| = \inf\{\|v_0 - u\| \mid u \in \mathsf{U}\}.$$

We claim that $v_0 - \hat{v}_0 \in \mathsf{U}^\perp$.

First we do a little computation. Let $v \in \mathsf{V}$, let $u \in \mathsf{U}\setminus\{0_{\mathsf{V}}\}$, and let $a \in \mathbb{F}$. Then we compute

$$\begin{aligned}\|v - au\|^2 &= \langle v - au, v - au\rangle \\ &= \|v\|^2 - a\langle u, v\rangle - \bar{a}\langle v, u\rangle + a\bar{a}\|u\|^2 \\ &= \|v\|^2 + \|u\|^2\Big(a\bar{a} - a\frac{\overline{\langle v, u\rangle}}{\|u\|^2} - \bar{a}\frac{\langle v, u\rangle}{\|u\|^2}\Big) \\ &= \|v\|^2 + \|u\|^2\Big(a - \frac{\langle v, u\rangle}{\|u\|^2}\Big)\Big(\bar{a} - \frac{\overline{\langle v, u\rangle}}{\|u\|^2}\Big) - \frac{|\langle v, u\rangle|^2}{\|u\|^2} \\ &= \|v\|^2 + \|u\|^2\Big|a - \frac{\langle v, u\rangle}{\|u\|^2}\Big|^2 - \frac{|\langle v, u\rangle|^2}{\|u\|^2}.\end{aligned}$$

As a function of a this quantity is minimised when $a = a_0 \triangleq \frac{\langle v,u\rangle}{\|u\|^2}$ and the minimum value of the function is

$$\|v\|^2 - \frac{|\langle v, u\rangle|^2}{\|u\|^2}.$$

Now apply this to $v = v_0 - \hat{v}_0$ to give

$$\|v_0 - \hat{v}_0 - a_0 u\| = \|v_0 - \hat{v}_0\| - \frac{|\langle v_0 - \hat{v}_0, u\rangle|^2}{\|u\|^2} \tag{7.5}$$

for every $u \in \mathsf{U} \setminus \{0_{\mathsf{V}}\}$. By definition of $\hat{v}_0$ we have

$$\|v_0 - \hat{v}_0 + a_0 u\|^2 \geq \|v_0 - \hat{v_0}\|^2,$$

and from this and (7.5) we have

$$\frac{|\langle v_0 - \hat{v}_0, u\rangle|^2}{\|u\|^2} = 0 \implies |\langle v_0 - \hat{v}_0, u\rangle|^2 = 0$$

for all $u \in \mathsf{U} \setminus \{0_{\mathsf{V}}\}$. Thus $v_0 - \hat{v}_0 \in \mathsf{U}^\perp$, as claimed above.

Therefore, for every $v \in \mathsf{V}$ we can write $v = (v - \hat{v}) - \hat{v}$ where $\hat{v} \in \mathsf{U}$ and $v - \hat{v} \in \mathsf{U}^\perp$. Since $\mathsf{U} \cap \mathsf{U}^\perp = \{0_{\mathsf{V}}\}$ by Proposition 7.1.13 we have $\mathsf{V} = \mathsf{U} \oplus \mathsf{U}^\perp$, giving the theorem. ■

As concerns Hilbert spaces, we have the following result.

7.1.20 Corollary (Closed subspaces and direct sum decompositions) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is an $\mathbb{F}$-Hilbert space, and if U is a closed subspace of V, then $\mathsf{V} = \mathsf{U} \oplus \mathsf{U}^\perp$.*

Proof By Proposition **??** closed subspaces of Hilbert spaces are complete. Thus the result follows from Theorem 7.1.19. ■

In finite dimensions the hypotheses of the theorem are always satisfied for any inner product.

7.1.21 Corollary (Orthogonal decompositions of finite-dimensional inner product spaces) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is a finite-dimensional $\mathbb{F}$-inner product space, and if U is a subspace of V, then $\mathsf{V} = \mathsf{U} \oplus \mathsf{U}^\perp$.*

Let us give an examples exploring the necessity of the hypotheses of the preceding results concerning direct sum decompositions.

7.1.22 Examples (Direct sum decomposition of inner product spaces)

1. The assumption in Theorem 7.1.19 that U is complete is essential. Indeed, consider the Hilbert space $(\ell^2(\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ with

$$\langle (a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}}\rangle = \sum_{j=1}^{\infty} a_j\bar{b}_j.$$

Take the subspace $\mathbb{F}_0^\infty$ which is not complete since its completion is $\ell^2(\mathbb{F})$ by Corollary 6.7.21. By Proposition 7.1.13 we have

$$(\mathbb{F}_0^\infty)^\perp = (\mathrm{cl}(\mathbb{F}_0^\infty))^\perp = \ell^2(\mathbb{F})^\perp = \{0_{\ell^2(\mathbb{F})}\}.$$

Thus we have $\ell^2(\mathbb{F}) \neq \mathbb{F}_0^\infty \oplus (\mathbb{F}_0^\infty)^\perp$.

2. Let us now consider the necessity that V be a Hilbert space in Corollary 7.1.20. We consider the incomplete inner product space $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$ and the subspace

$$\mathsf{U} = \Big\{(a_j)_{j\in\mathbb{Z}_{>0}} \;\Big|\; \sum_{j=1}^{\infty}\frac{a_j}{j} = 0\Big\}.$$

We leave to the reader the elementary verification that U is a proper subspace of $\mathbb{F}_0^\infty$.

Let us verify that U is closed. Let $((a_{jl})_{j\in\mathbb{Z}_{>0}})_{l\in\mathbb{Z}_{>0}}$ be a sequence in U converging to $(a_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{F}_0^\infty$. Fix $j\in\mathbb{Z}_{>0}$ and let $\epsilon\in\mathbb{R}_{>0}$. Choose $N\in\mathbb{Z}_{>0}$ sufficiently large that

$$\|(a_j)_{j\in\mathbb{Z}_{>0}} - (a_{jl})_{j\in\mathbb{Z}_{>0}}\| < \epsilon$$

for $l \geq N$. Then, for $l\geq N$,

$$|a_j - a_{jl}|^2 \leq \sum_{k=1}^{\infty}|a_k - a_{kl}|^2 = \|(a_k)_{k\in\mathbb{Z}_{>0}} - (a_{kl})_{k\in\mathbb{Z}_{>0}}\|^2 < \epsilon^2.$$

That is to say, $\lim_{l\to\infty} a_{jl} = a_j$ for each $j\in\mathbb{Z}_{>0}$. Define

$$b_{nl} = \sum_{j=1}^{n}\frac{a_{jl}}{j}.$$

We claim that the double sequence $(b_{nl})_{n,l\in\mathbb{Z}}$ converges to zero. Since $(a_j)_{j\in\mathbb{Z}_{>0}}\in\mathbb{F}_0^\infty$ there exists $N_1\in\mathbb{Z}_{>0}$ such that $a_j = 0$ for $j > N_1$. Now let $\epsilon\in\mathbb{R}_{>0}$ and let $N_2\in\mathbb{Z}_{>0}$ be sufficiently large that

$$\|(a_j)_{j\in\mathbb{Z}_{>0}} - (a_{jl})_{j\in\mathbb{Z}_{>0}}\| < \frac{\epsilon}{M}$$

for $l \geq N_2$, where

$$M \triangleq \sum_{j=1}^{\infty}\frac{1}{j^2},$$

this series being summable by Example 2.4.2–**??**. Then, using the Cauchy–Bunyakovsky–Schwarz inequality, for $l, n \geq \max\{N_1, N_2\}$,

$$\begin{aligned}|b_{nl}| = \Big|\sum_{j=1}^{n} \frac{a_{jl}}{j}\Big| &\leq \Big|\sum_{j=1}^{n} \frac{a_{jl} - a_j}{j}\Big| + \Big|\sum_{j=1}^{n} \frac{a_j}{j}\Big| \\ &\leq \Big(\sum_{j=1}^{n} |a_j - a_{jl}|^2\Big)^{1/2} \Big(\sum_{j=1}^{n} \frac{1}{j^2}\Big)^{1/2} \\ &\leq \Big(\sum_{j=1}^{\infty} |a_j - a_{jl}|^2\Big)^{1/2} \Big(\sum_{j=1}^{\infty} \frac{1}{j^2}\Big)^{1/2} < \epsilon,\end{aligned}$$

as desired. Then we have

$$\sum_{j=1}^{\infty} \frac{a_j}{j} = \sum_{j=1}^{\infty} \lim_{l\to\infty} \frac{a_{jl}}{j} = \lim_{n\to\infty} \lim_{l\to\infty} b_{nl} = 0,$$

using Proposition 2.3.21. Thus we indeed have $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{U}$ and so U is closed. Now let us show that $\mathsf{U}^\perp = \{0_{\mathbb{F}_0^\infty}\}$. Let $(a_j)_{j\in\mathbb{Z}_{>0}} \in \mathsf{U}^\perp$ and let $N \in \mathbb{Z}_{>0}$ be such that $a_j = 0$ for $j > N$. Then define $(b_{jl})_{j\in\mathbb{Z}_{>0}} \in \mathsf{U}$, $l \in \{1, \dots, N+1\}$, by

$$b_{jl} = \begin{cases} -l, & j = l, \\ N+1, & j = N+1, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\sum_{j=1}^{\infty} \frac{b_{jl}}{j} = -\frac{l}{l} + \frac{N+1}{N+1} = 0,$$

so $(b_{jl})_{j\in\mathbb{Z}_{>0}}$ is indeed in U for each $l \in \{1, \dots, N+1\}$. Moreover, for each $l \in \{1, \dots, N\}$,

$$0 = \langle (a_j)_{j\in\mathbb{Z}_{>0}}, (b_{jl})_{j\in\mathbb{Z}_{>0}} \rangle = -la_l,$$

and so $a_l = 0$ for $l \in \{1, \dots, N\}$. Thus $\mathsf{U}^\perp = \{0_{\mathbb{F}_0^\infty}\}$ as claimed. •

The preceding examples suggest that there is some sort of relationship between completeness of inner product spaces and properties of closed subspaces. Let us clarify this with the following result.

7.1.23 Theorem (Subspace characterisations of completeness of inner product spaces) *For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and for a $\mathbb{F}$-inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$, the following statements are equivalent:*

(i) V is a Hilbert space;

(ii) for every closed subspace U of V it holds that $\mathsf{V} = \mathsf{U} \oplus \mathsf{U}^\perp$;

(iii) for every closed subspace U of V it holds that $\mathsf{U} = (\mathsf{U}^\perp)^\perp$;

(iv) for every proper closed subspace U of V it holds that $\mathsf{U}^\perp \neq \{0_{\mathsf{V}}\}$.

Proof (i) $\Longrightarrow$ (ii) This is Corollary 7.1.20.

(ii) $\Longrightarrow$ (iii) By Proposition 7.1.13 we have $\mathsf{U} \subseteq (\mathsf{U}^\perp)^\perp$. Now let $v \in (\mathsf{U}^\perp)^\perp$ and write $v = v_1 + v_2$ for $v_1 \in \mathsf{U}$ and $v_2 \in \mathsf{U}^\perp$. Then $v_2 = v - v_1 \in (\mathsf{U}^\perp)^\perp$ since $v \in (\mathsf{U}^\perp)^\perp$ and $v_1 \in \mathsf{U} \subseteq (\mathsf{U}^\perp)^\perp$. But this means that $v_2 \in \mathsf{U}^\perp \cap (\mathsf{U}^\perp)^\perp = \{0_{\mathsf{V}}\}$ and so $v = v_1 \in \mathsf{U}$.

(iii) $\Longrightarrow$ (iv) Let U be a subspace of V for which $\mathsf{U}^\perp = \{0_{\mathsf{V}}\}$. By assumption, $\mathsf{U} = \{0_{\mathsf{V}}\}^\perp = \mathsf{V}$. Thus U is not proper.

(iv) $\Longrightarrow$ (i) Let $\overline{\mathsf{V}}$ be a completion of V and regard V as a subspace of $\overline{\mathsf{V}}$. Let $\bar{v} \in \overline{\mathsf{V}}$. If $\bar{v} = 0_{\mathsf{V}}$ then $\bar{v} \in \mathsf{V}$. So suppose that $\bar{v} \neq 0_{\mathsf{V}}$. Define $f_{\bar{v}} \colon \mathsf{V} \to \mathbb{F}$ by $f_{\bar{v}}(u) = \langle u, \bar{v} \rangle$ noting that $f_{\bar{v}}$ is continuous by Proposition 7.2.1. Thus $\ker(f_{\bar{v}})$ is closed by Theorem 6.5.2, being the preimage of the closed set $\{0_{\mathsf{V}}\}$. We claim that $\ker(f_{\bar{v}})$ is a proper subspace. To see this, suppose that $\ker(f_{\bar{v}}) = \mathsf{V}$ and let $(v_j)_{j \in \mathbb{Z}_{>0}}$ be a sequence in V converging to $\bar{v}$. Then, by Theorem 6.5.2 and Proposition 7.2.1, we have

$$\langle \bar{v}, \bar{v} \rangle = \langle \lim_{j\to\infty} v_j, \bar{v} \rangle = \lim_{j\to\infty} \langle v_j, \bar{v} \rangle = 0,$$

contradicting the definiteness of the inner product. Thus we have $\ker(f_{\bar{v}}) \subset \mathsf{V}$. By assumption there exists $v' \in \ker(f_{\bar{v}})^\perp$ such that $\|v\| = 1$. One can verify, cf. the proof of Theorem 7.2.2 below, that if we take $v = \overline{f_v(v')}v'$ then $\langle u, \bar{v} \rangle = \langle u, v \rangle$ for every $u \in \mathsf{V}$. Thus $\langle u, \bar{v} - v \rangle = 0$ for every $u \in \mathsf{V}$ and so $\bar{v} = v$. Thus $\overline{\mathsf{V}} = \mathsf{V}$. ■

For other conditions equivalent to completeness we refer to Theorems 7.2.4 and 7.3.10.

7.1.5 Minimising distance to a set

One of the very interesting and useful features of inner product spaces is that they allow one to solve certain sorts of problems. In this section we consider the following problem.

7.1.24 Problem (Distance minimisation problem) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a $\mathbb{F}$-normed vector space. For $v_0 \in \mathsf{V}$ and for a subset $S \subseteq \mathsf{V}$ do the following:

(i) determine $\mathrm{dist}(v_0, S) \triangleq \inf\{\|v_0 - v\| \mid v \in S\}$;

(ii) ascertain whether there exists $\hat{v}_0 \in S$ such that $\|v_0 - \hat{v}_0\| = \mathrm{dist}(v_0, S)$. •

In general, the previous problem is too difficult to be approachable. There are a couple of reasons for this. First of all, by stating the problem for arbitrary subsets the problem is simply unreasonable. One really must place some additional structure on the set S. Below we will consider the case when S is convex. However, even if one restricts the set S to be something "reasonable," the problem can still be too difficult to solve. One of the reasons this may be so is that general norms are difficult to understand. The reader can explore this a little in the finite-dimensional situation in Exercise 7.1.14. However, if one restricts the norm to come from an inner product it turns out that it is possible to characterise the solutions to some distance minimisation problems in a useful way. Thus we restrict our attention in this section to the distance minimisation problem for inner product spaces.

The most accessible sufficiently interesting result concerns the minimisation of the distance from a point to a convex set. We dealt with convexity in $\mathbb{R}^n$ in detail in

Section **??** and in general vector spaces in Chapter **??**. Here we simply recall that a convex subset of a $\mathbb{F}$-vector space V is a subset C for which

$$u, v \in C \quad \Longrightarrow \quad \{(1-s)u + sv \mid s \in [0,1]\} \subseteq C.$$

We then have the following result which gives a case where the distance minimisation problem possesses a unique solution.

7.1.25 Theorem (Minimisation of distance to convex subsets) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*-inner product space, and let* $\mathrm{v}_0 \in \mathsf{V}$*. If* $\mathrm{C} \subseteq \mathsf{V}$ *is a complete convex set then there exists a unique vector* $\hat{\mathrm{v}}_0 \in \mathrm{C}$ *for which*

$$\|\mathrm{v}_0 - \hat{\mathrm{v}}_0\| = \mathrm{dist}(\mathrm{v}_0, \mathrm{C}).$$

Proof Denote $m = \mathrm{dist}(v, C)$ and let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in C such that $\|v_0 - v_j\|^2 < m^2 + \frac{1}{j}$. We claim that the set

$$\{v_0\} + C = \{v_0 + v \mid v \in C\}$$

is convex. Indeed, if $v_0 + v_1, v_0 + v_2 \in \{v_0\} + C$ for $v_1, v_2 \in C$ and if $s \in [0,1]$ then

$$(1-s)(v_0 + v_1) + s(v_0 + v_2) = v_0 + (1-s)v_1 + sv_2 \in \{v_0\} + C.$$

Now, since $\{v_0\} + C$ is convex, for each $j, k \in \mathbb{Z}_{>0}$ we have $\left\|\frac{1}{2}((v_0 + v_j) + (v_0 + v_k))\right\|^2 \geq m^2$. Now let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that $\frac{4}{N} < \epsilon^2$. For $j, k \geq N$, using the parallelogram law we then have

$$\begin{aligned}\|v_j - v_k\|^2 &= \|(v_0 - v_j) - (v_0 - v_k)\|^2 \\ &= 2\|v_0 - v_j\|^2 + 2\|v_0 - v_k\|^2 - 4\left\|\tfrac{1}{2}((v_0 + v_j) + (v_0 + v_k))\right\|^2 \\ &< 2m^2 + \tfrac{2}{j} + 2m^2 + \tfrac{2}{k} - 4m^2 < \tfrac{4}{N} < \epsilon^2.\end{aligned}$$

Thus $\|v_j - v_k\| < \epsilon$ for $j, k \geq N$ and so $(v_j)_{j\in\mathbb{Z}_{>0}}$ is a Cauchy sequence. Since C is complete there exists $\hat{v}_0 \in C$ such that $(v_j)_{j\in\mathbb{Z}_{>0}}$ converges to $\hat{v}_0$. This gives the existence part of the lemma.

If $\hat{u}_0 \in C$ has the property that $\|v_0 - \hat{u}_0\| = m$ then, using the parallelogram law,

$$\begin{aligned}\|\hat{u}_0 - \hat{v}_0\|^2 = 2\|v_0 - \hat{u}_0\|^2 &+ 2\|v_0 - \hat{v}_0\|^2 \\ &- 4\left\|\tfrac{1}{2}((v_0 + \hat{u}_0) + (v_0 + \hat{v}_0))\right\| \leq 2m^2 + 2m^2 - 4m^2 = 0.\end{aligned}$$

Thus $\|\hat{u}_0 - \hat{v}_0\| = 0$ and so $\hat{u}_0 = \hat{v}_0$. ■

Since a subspace of a vector space is obviously convex we can immediately apply the preceding result to the case when C is a subspace. For subspaces, however, there is more that can be said about the character of the points that solve the distance minimisation problem: they are orthogonal to the subspace.

7.1.26 Theorem (Minimisation of distance to subspaces) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*-inner product with* $\mathrm{v}_0 \in \mathsf{V}$ *and* $\mathsf{U} \subseteq \mathsf{V}$ *a subspace. Then* $\hat{\mathrm{v}}_0 \in \mathsf{U}$ *satisfies*

$$\|\mathrm{v}_0 - \hat{\mathrm{v}}_0\| = \mathrm{dist}(\mathrm{v}_0, \mathsf{U}) \tag{7.6}$$

if and only if $\mathrm{v}_0 - \hat{\mathrm{v}}_0 \in \mathsf{U}^{\perp}$*. Furthermore, if* U *is complete then there exists a unique vector* $\hat{\mathrm{v}}_0 \in \mathsf{U}$ *such that (7.6) holds.*

Proof First suppose that $v_0 - \hat{v}_0 \in \mathsf{U}^{\perp}$. Then, since $\hat{v}_0 - u \in \mathsf{U}$ for any $u \in \mathsf{U}$, $v_0 - \hat{v}_0$ and $\hat{v}_0 - u$ are orthogonal. The Pythagorean identity (Exercise 7.1.12) then gives

$$\|v_0 - u\|^2 = \|v_0 - \hat{v}_0\|^2 + \|\hat{v}_0 - u\|^2$$

for any $u \in \mathsf{U}$. From this we conclude that $\|v_0 - \hat{v}_0\|^2 \leq \|v_0 - u\|^2$ for every $u \in \mathsf{U}$. This exactly means that $\hat{v}_0$ satisfies (7.6).

Now suppose that $\hat{v}_0$ satisfies (7.6). Let $\alpha \in \mathbb{F} \setminus \{0\}$ and define $f_\alpha \colon \mathsf{U} \to \mathsf{U}$ by $f_\alpha(u) = \hat{v}_0 + \alpha(u - \hat{v}_0)$. Since $\hat{v}_0$ satisfies (7.6) we have

$$\begin{aligned}\|v_0 - \hat{v}_0\|^2 &\leq \|v_0 - f_\alpha(u)\|^2\\ &= \|(v_0 - \hat{v}_0) - \alpha(u - \hat{v}_0)\|^2\\ &= \|v_0 - \hat{v}_0\|^2 + |\alpha|^2\|u - \hat{v}_0\| - \alpha\langle u - \hat{v}_0, v_0 - \hat{v}_0\rangle - \bar{\alpha}\langle v_0 - \hat{v}_0, u - \hat{v}_0\rangle.\end{aligned}$$

From this we conclude that

$$\alpha\langle u - \hat{v}_0, v_0 - \hat{v}_0\rangle + \bar{\alpha}\overline{\langle u - \hat{v}_0, v_0 - \hat{v}_0\rangle} \leq |\alpha|^2\|u - \hat{v}_0\| \tag{7.7}$$

for every $u \in \mathsf{U}$. Now we write $\alpha = |\alpha|e^{\mathrm{i}\theta}$ for $\theta \in (-\pi, \pi]$. If $\mathbb{F} = \mathbb{R}$ we restrict to $\theta \in \{0, \pi\}$. Now divide (7.7) by $|\alpha|$ and take the limit as $|\alpha| \to 0$. Also note that

$$\{u - \hat{v}_0 \mid u \in \mathsf{U}\} = \mathsf{U}.$$

Putting this all together gives

$$e^{\mathrm{i}\theta}\langle u, v_0 - \hat{v}_0\rangle + e^{-\mathrm{i}\theta}\overline{\langle u, v_0 - \hat{v}_0\rangle} \leq 0,$$

which again holds for all $u \in \mathsf{U}$ and $\theta \in (-\pi, \pi]$. Taking $\theta = 0$ gives

$$2\,\mathrm{Re}(\langle u, v_0 - \hat{v}_0\rangle) \leq 0,$$

and taking $\theta = \pi$ gives

$$-2\,\mathrm{Re}(\langle u, v_0 - \hat{v}_0\rangle) \leq 0$$

for all $u \in \mathsf{U}$. From this we conclude that $\mathrm{Re}(\langle u, v_0 - \hat{v}_0\rangle) = 0$ for all $u \in \mathsf{U}$. A similar argument, using $\theta = \frac{\pi}{2}$ and $\theta = -\frac{\pi}{2}$, gives $\mathrm{Im}(\langle u, v_0 - \hat{v}_0\rangle) = 0$. Thus $v_0 - \hat{v}_0 \in \mathsf{U}^{\perp}$, as desired.

The final assertion of the theorem follows directly from Theorem 7.1.25. ■

The preceding result is insightful as it gives us a concrete description of the set of points that minimise the distance from a vector v_0 to a subspace U. This description will be important for us in Section 7.3.4 subsequently for applications of the ideas in Section 7.3.4. You will observe that the most difficult part of Theorem 7.1.26 is showing that the set of points minimising the distance is nonempty, and in fact contains a single point, at least when U is complete. In finite-dimensions, these issues are not so complicated, as can be seen in Exercise 7.1.15.

missing stuff

7.1.6 Norms

Theorem 7.1.9 was proved by John von Neumann.

Example 7.1.22–2 is taken from [**SG:74b**], as are the characterisations of completeness in Theorem 7.1.23.

Exercises

7.1.1 Prove Proposition 7.1.2.

7.1.2 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a $\mathbb{F}$-(semi-)inner product space. For a subspace $\mathsf{U} \subseteq \mathsf{V}$, define a map from $\mathsf{U}\times\mathsf{U}$ to $\mathbb{F}$ by $(u_1, u_2) \mapsto \langle u_1, u_2\rangle \triangleq \langle u_1, u_2\rangle_{\mathsf{U}}$. Show that $(\mathsf{U}, \langle\cdot,\cdot\rangle_{\mathsf{U}})$ is an $\mathbb{F}$-(semi-)inner product space.

7.1.3 Show that a $\mathbb{C}$-inner product space is always a $\mathbb{R}$-inner product space, using the fact that a $\mathbb{C}$-vector space is always a $\mathbb{R}$-vector space.

7.1.4 Answer the following three questions.

(a) Show that the norm defined by an inner product satisfies the parallelogram law.

(b) Show that the norm defined in Example 6.1.3–4 does not come from an inner product.

(c) Give an interpretation of the parallelogram law in $\mathbb{R}^2$ with the standard inner product.

7.1.5 Show using the parallelogram law that the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ on $\mathbb{F}^n$ are not derived from an inner product if $n \geq 2$.

7.1.6 Show explicitly (i.e., as is done in Example 6.3.1–1) that $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$ is not complete.

7.1.7 Show explicitly (i.e., as is done in Example 6.3.1–2) that $(\mathsf{C}^0([a,b],\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ is not complete.

7.1.8 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Show that the following two assertions are equivalent:

(i) there exists a (semi-)inner product $\langle\cdot,\cdot\rangle$ on V such that $\|v\| = \sqrt{\langle v, v\rangle}$ for every $v \in \mathsf{V}$;

(ii) the expression

$$\|u+v+w\|^2 + \|u+v-w\|^2 - \|u-v-w\|^2 - \|u-v+w\|^2$$

is independent of w.

Hint: *Use Theorem 7.1.9.*

7.1.9 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \|\cdot\|)$ be a (semi)normed $\mathbb{F}$-vector space. Show that the following two assertions are equivalent:

(i) there exists a (semi-)inner product $\|\cdot\|\cdot$ on V such that $\|v\| = \sqrt{\langle v, v\rangle}$ for every $v \in \mathsf{V}$;

(ii) the function $s \mapsto \|u + sv\|^2$ is a polynomial function of degree 2 for every $u, v \in \mathsf{V}$.

Hint: *Use Theorem 7.1.9.*

7.1.10 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space. Show that if subsets $A, B \subseteq \mathsf{V}$ are orthogonal then so too are the subsets $\operatorname{span}_{\mathbb{F}}(A)$ and $\operatorname{span}_{\mathbb{F}}(B)$.

7.1.11 Prove parts (i), (ii), and (iii) of Proposition 7.1.13.

7.1.12 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a $\mathbb{F}$-semi-inner product space.

(a) Prove the ***Pythagorean identity***:

$$\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2$$

if v_1 and v_2 are orthogonal.

(b) Show that if $\mathbb{F} = \mathbb{R}$ then the Pythagorean identity for v_1 and v_2 implies that v_1 and v_2 are orthogonal.

(c) Give an example showing that the assertion in part (b) is generally false if $\mathbb{F} = \mathbb{C}$.

7.1.13 For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, for an $\mathbb{F}$-inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$, and for a subspace $\mathsf{U} \subseteq \mathsf{V}$, answer the following two questions.

(a) Show that $\mathsf{U} \cap \mathsf{U}^{\perp} = \{0\}$.

(b) Show that if U is closed then for every $v \in \mathsf{V}$ there exists unique vectors $u_1 \in \mathsf{U}$ and $u_2 \in \mathsf{U}^{\perp}$ so that $v = u_1 + u_2$.

7.1.14 Consider the Banach space $(\mathbb{R}^2, \|\cdot\|_2)$ of Example 6.1.3–2 and the Banach space $(\mathbb{R}^2, \|\cdot\|_\infty)$ of Example 6.1.3–4. For each of these norms, and for the subsets S and the points $\boldsymbol{v}_0$ given below, determine $\operatorname{dist}(\boldsymbol{v}_0, S)$ and determine the set of points $\hat{\boldsymbol{v}}_0 \in S$ such that $\|\boldsymbol{v}_0 - \hat{\boldsymbol{v}}_0\| = \operatorname{dist}(\boldsymbol{v}_0, S)$.

(a) $\boldsymbol{v}_0 = (0,1)$ and $S = \{(v_1, 0) \mid v_1 \in [-1,1]\}$.

(b) $\boldsymbol{v}_0 = (0,1)$ and $S = \operatorname{span}_{\mathbb{R}}((1,0))$.

(c) $\boldsymbol{v}_0 = (0,1)$ and $S = \operatorname{span}_{\mathbb{R}}((1,1))$.

(d) $\boldsymbol{v}_0 = (0,0)$ and $S = \{(v_1, v_2) \mid v_1^2 + v_2^2 \geq 1\}$.

(e) $\boldsymbol{v}_0 = (0,0)$ and $S = \{(v_1, v_2) \mid v_1^2 + v_2^2 > 1\}$.

In the next exercise you will prove Theorem 7.1.26 when V is finite-dimensional. As you will see, it is possible to be somewhat more concrete in this case, making you appreciate that there is something real happening in the proof of Theorem 7.1.26.

7.1.15 Let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space, and let $v_0 \in \mathsf{V}$ with $\mathsf{U} \subseteq \mathsf{V}$ a subspace. Provide a proof of Theorem 7.1.26 in this case along the following lines.

1. Argue that the result is trivial unless $v_0 \notin \mathsf{U}$. Thus assume this for the remainder of the proof.
2. For a subspace $\mathsf{U} \subseteq \mathsf{V}$ let $\{u_1, \dots, u_m\}$ be an orthonormal basis for U. Can this always be done?
3. Extend the basis from the previous part of the question to an orthonormal basis $\{v_1 = u_1, \dots, v_m = u_m, v_{m+1}, \dots, v_n\}$ for V. Can this always be done?

4. As a function on U, use the above basis to explicitly write down the function defining the distance from U to v_0.
5. Show that the unique point in U that minimises the distance function is

$$\hat{v}_0 = \sum_{j=1}^{m} \langle v, u_j \rangle u_j.$$

Section 7.2

Continuous maps between inner product spaces

Inner product spaces, being normed vector spaces, are of course subject to all the definitions and results concerning maps between normed vector spaces as stated in Section 6.5. We shall take all of these definitions and results for granted, and instead emphasise the things that are distinctive for inner product spaces.

Do I need to read this section? The results in this section complement those of Section 6.5, and so should be absorbed if one is in the business of understanding continuous maps between infinite-dimensional spaces. •

7.2.1 The dual of an inner product space

Much of the special character of inner product spaces, as opposed to more general normed vector spaces, is reflected in the structure of the topological dual of an inner product space. In order to understand this it is useful to first record some elementary properties of inner products.

7.2.1 Proposition (Continuity properties of operations in an inner product space) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an inner product space. Then the following maps are uniformly continuous:*

(i) $\mathsf{V} \ni \mathrm{v} \mapsto \langle \mathrm{v}, \mathrm{v}_0\rangle \in \mathbb{F}$ *for* $\mathrm{v}_0 \in \mathsf{V}$;

(ii) $\mathsf{V} \ni \mathrm{v} \mapsto \langle \mathrm{v}_0, \mathrm{v}\rangle \in \mathbb{F}$ *for* $\mathrm{v}_0 \in \mathsf{V}$;

(iii) $\mathbb{F} \ni \mathrm{a} \mapsto \langle \mathrm{a}\mathrm{v}_1, \mathrm{v}_2\rangle \in \mathbb{F}$ *for* $\mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V}$;

(iv) $\mathbb{F} \ni \mathrm{a} \mapsto \langle \mathrm{v}_1, \mathrm{a}\mathrm{v}_2\rangle \in \mathbb{F}$ *for* $\mathrm{v}_1, \mathrm{v}_2 \in \mathsf{V}$.

Proof (i) If $v_0 = 0_{\mathsf{V}}$ the assertion is clearly true as the map is the constant map with value zero. Thus consider $v_0 \neq 0_{\mathsf{V}}$. Let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \frac{\epsilon}{\|v_0\|}$. Then, using the Cauchy–Bunyakovsky–Schwarz inequality,

$$|\langle v_1, v_0\rangle - \langle v_2, v_0\rangle| = |\langle v_1 - v_2, v_0\rangle| \leq \|v_1 - v_2\|\|v_0\| \leq \epsilon$$

for $\|v_1 - v_2\| < \delta$.

(ii) Conjugation $a \mapsto \bar{a}$ is clearly uniformly continuous. Therefore, $v \mapsto \langle v_0, v\rangle = \overline{\langle v, v_0\rangle}$ is uniformly continuous, being a composition of uniformly continuous maps.

(iii) If $\langle v_1, v_2\rangle = 0$ then clearly the given map is continuous since it is the constant map with value zero. So suppose that $\langle v_1, v_2\rangle$ is nonzero. Let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \frac{\epsilon}{|\langle v_1, v_2\rangle|}$. Then

$$|\langle a_1 v_1, v_2\rangle - \langle a_2 v_1, v_2\rangle| = |\langle (a_1 - a_2) v_1, v_2\rangle| = |a_1 - a_2||\langle v_1, v_2\rangle| \leq \epsilon$$

for $|a_1 - a_2| < \delta$.

(iv) This follows from part (iii) as part (ii) follows from (i). ■

The central result concerning the dual of an inner product space is then the following.

7.2.2 Theorem (Riesz Representation Theorem) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be a Hilbert space with topological dual* V^*. *If* $\alpha \in \mathsf{V}^*$ *then there exists a unique* $\mathrm{v}_\alpha \in \mathsf{V}$ *such that* $\langle \mathrm{u}, \mathrm{v}_\alpha \rangle = \alpha(\mathrm{u})$ *for every* $\mathrm{u} \in \mathsf{V}$.

Proof If $\alpha = 0$ then we can take $v_\alpha = 0$. So let $\alpha \in \mathsf{V}^* \setminus \{0\}$. We claim that $\ker(\alpha)$ is a closed subspace of V. It is certainly a subspace. To show that it is closed, let $(v_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\ker(\alpha)$ converging to $v_0 \in \mathsf{V}$. Then, by continuity of α and Theorem 6.5.2 we have

$$\alpha(v_0) = \alpha\Big(\lim_{j\to\infty} v_j\Big) = \lim_{j\to\infty} \alpha(v_j) = 0.$$

Thus $v_0 \in \ker(\alpha)$ and so $\ker(\alpha)$ is closed by Proposition 6.6.8. Since $\alpha \neq 0$, $\ker(\alpha) \neq \mathsf{V}$. By Theorem 7.1.19, since $\ker(\alpha)$ is closed we can choose a nonzero vector $v_0 \in \ker(\alpha)^\perp$, supposing this vector to further have length 1. We claim that we can take $v_\alpha = \bar{\alpha}(v_0)v_0$, where $\bar{\alpha}\colon \mathsf{V} \to \mathbb{F}$ is defined by $\bar{\alpha}(v) = \overline{\alpha(v)}$. Indeed note that for $u \in \mathsf{V}$ the vector $\alpha(u)v_0 - \alpha(v_0)u$ is in $\ker(\alpha)$. Therefore

$$0 = \langle \alpha(u)v_0 - \alpha(v_0)u, v_0\rangle = \alpha(u) - \alpha(v_0)\langle u, v_0\rangle.$$

Thus

$$\alpha(u) = \langle u, \bar{\alpha}(v_0)v_0\rangle = \langle u, v_\alpha\rangle.$$

Thus v_α as defined meets the desired criterion. Let us show that this is the only vector satisfying the conditions of the theorem. Suppose that $v_1, v_2 \in \mathsf{V}$ have the property that $\alpha(u) = \langle u, v_1\rangle = \langle u, v_2\rangle$ for all $u \in \mathsf{V}$. Then $\langle u, v_1 - v_2\rangle = 0$ for all $u \in \mathsf{V}$. In particular, taking $u = v_1 - v_2$ we have $\|v_1 - v_2\|^2 = 0$, giving $v_1 = v_2$. ■

The assumption that V is a Hilbert space is essential as the following example shows.

7.2.3 Example (The dual of an incomplete inner product space) Let us consider the $\mathbb{F}$-inner product space $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$ where, we recall, that

$$\langle (a_j)_{j\in\mathbb{Z}_{>0}}, (b_j)_{j\in\mathbb{Z}_{>0}}\rangle = \sum_{j=1}^\infty a_j\bar{b}_j;$$

the sum is finite. Recall from Proposition **??** that $(\mathbb{F}_0^\infty)' = \mathbb{F}^\infty$ and so $(\mathbb{F}_0^\infty)^*$ is a subspace of $\mathbb{F}^\infty$. Define $\boldsymbol{\alpha} \in \mathbb{F}^\infty$ by $\boldsymbol{\alpha}(j) = \frac{1}{j}$ for each $j \in \mathbb{Z}_{>0}$. By Example 2.4.2–**??** note that

$$(\tfrac{1}{j})_{j\in\mathbb{Z}_{>0}} \in \ell^2(\mathbb{F}) \subseteq \mathbb{F}^\infty \quad\implies\quad M^2 \triangleq \sum_{j=1}^\infty \frac{1}{j^2} < \infty.$$

(In fact, $M^2 = \frac{\pi^2}{6}$ but this precise number is not important for us, only that it is finite.)

We claim that $\boldsymbol{\alpha}$ is a continuous linear function on $\mathbb{F}_0^\infty$. Indeed, let $\epsilon \in \mathbb{R}_{>0}$ and take $\delta = \frac{\epsilon}{M}$. Let $\boldsymbol{a} = (a_j)_{j\in\mathbb{Z}_{>0}}, \boldsymbol{b} = (b_j)_{j\in\mathbb{Z}_{>0}}$ be such that

$$\|(a_j)_{j\in\mathbb{Z}_{>0}} - (b_j)_{j\in\mathbb{Z}_{>0}}\|_2 < \delta.$$

Then, using the Cauchy–Bunyakovsky–Schwarz inequality,

$$|\boldsymbol{\alpha}(\boldsymbol{a}) - \boldsymbol{\alpha}(\boldsymbol{b})| = |\boldsymbol{\alpha}(\boldsymbol{a} - \boldsymbol{b})| = \Big|\sum_{j=1}^{\infty} \frac{a_j - b_j}{j^2}\Big| \le \Big(\sum_{j=1}^{\infty} |a_j - b_j|^2\Big)^{1/2} \Big(\sum_{j=1}^{\infty} \frac{1}{j^2}\Big)^{1/2} < \epsilon.$$

Thus $\boldsymbol{\alpha}$ is indeed continuous.

We next claim that there exists no $\boldsymbol{f}_{\boldsymbol{\alpha}} \in \mathbb{F}_0^\infty$ such that $\langle \boldsymbol{f}_{\boldsymbol{\alpha}}, \boldsymbol{a}\rangle = \boldsymbol{\alpha}(\boldsymbol{a})$ for every $\boldsymbol{a} \in \mathbb{F}_0^\infty$. To see this, let $(\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ be the standard basis for $\mathbb{F}_0^\infty$ so that $\boldsymbol{e}_j(k) = 1$ for $j = k$ and 0 otherwise. Then, if $\boldsymbol{f}_{\boldsymbol{\alpha}} \in \mathbb{F}_0^\infty$ we have $\langle \boldsymbol{f}_{\boldsymbol{\alpha}}, \boldsymbol{e}_j\rangle = \boldsymbol{f}_{\boldsymbol{\alpha}}(j)$. Also, $\boldsymbol{\alpha}(\boldsymbol{e}_j) = \frac{1}{j}$ for each $j \in \mathbb{Z}_{>0}$. Thus if $\boldsymbol{f}_{\boldsymbol{\alpha}} \in \mathbb{F}^\infty$ has the property that $\langle \boldsymbol{f}_{\boldsymbol{\alpha}}, \boldsymbol{e}_j\rangle = \boldsymbol{\alpha}(\boldsymbol{e}_j)$ for every $j \in \mathbb{Z}_{>0}$ then it follows that $\boldsymbol{f}_{\boldsymbol{\alpha}}(j) = \frac{1}{j}$ for each $j \in \mathbb{Z}_{>0}$. But this means that $\boldsymbol{f}_{\boldsymbol{\alpha}}(j) \notin \mathbb{F}_0^\infty$. •

The preceding example is, actually, representative of the general situation in the sense of the following result which states that the assumption that V be a Hilbert space is essential in the Riesz Representation Theorem.

7.2.4 Theorem (The Riesz Representation Theorem does not hold for non-Hilbert spaces) *For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and for a $\mathbb{F}$-inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$, the following statements are equivalent:*

(i) V is a Hilbert space;

(ii) for every $\alpha \in \mathsf{V}^*$ there exists $\mathrm{v}_\alpha \in \mathsf{V}$ such that $\langle \mathrm{u}, \mathrm{v}_\alpha\rangle = \alpha(\mathrm{u})$ for every $\mathrm{u} \in \mathsf{V}$.

Proof That (i) $\Longrightarrow$ (ii) is simply Theorem 7.2.2, so we need only prove the converse. Thus we let $\overline{\mathsf{V}}$ be a completion of V, let $\bar{v} \in \overline{\mathsf{V}}$, and define $f_{\bar{v}} \colon \mathsf{V} \to \mathbb{F}$ by $f_{\bar{v}}(u) = \langle u, \bar{v}\rangle$. By Proposition 7.2.1 it follows that $f_{\bar{v}}$ is continuous. By assumption there exists $v \in \mathsf{V}$ such that $\langle u, v\rangle = f_{\bar{v}}(u) = \langle u, \bar{v}\rangle$ for every $u \in \mathsf{V}$. Thus $v = \bar{v}$ and so $\overline{\mathsf{V}} = \mathsf{V}$. ■

Let us examine a consequence of the Riesz Representation Theorem.

7.2.5 Corollary (The dual of a Hilbert space) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-Hilbert space. Then the map $\alpha \mapsto \mathrm{v}_\alpha$ from V^* to V is an isomorphism of $\mathbb{R}$-normed vector spaces that further satisfies $\mathrm{v}_{a\alpha} = \bar{a}\mathrm{v}_\alpha$.*

Proof According to the proof of Theorem 7.2.2 we have $v_\alpha = \bar{\alpha}(v_0)v_0$, where $\bar{\alpha} \in \mathsf{V}^*$ is defined by $\bar{\alpha}(v) = \overline{\alpha(v)}$ and where v_0 is a fixed vector of unit length in $\ker(\alpha)^\perp$. The conclusions of the corollary are directly verified. ■

Note that V^* are V are not isomorphic as $\mathbb{F}$-vector spaces in the case when $\mathbb{F} = \mathbb{C}$. Sometimes the property of a linear map $\mathsf{L} \colon \mathsf{U} \to \mathsf{V}$ that

1. $\mathsf{L}(u_1 + u_2) = \mathsf{L}(u_1) + \mathsf{L}(u_2)$, $u_1, u_2 \in \mathsf{U}$, and
2. $\mathsf{L}(au) = \bar{a}\mathsf{L}(u)$, $a \in \mathbb{F}$, $u \in \mathsf{U}$,

is called ***conjugate linearity*** and agree with the property of linearity if and only if $\mathbb{F} = \mathbb{R}$.

Let us examine the Riesz Representation Theorem in a few special cases.

7.2.6 Examples (Riesz Representation Theorem)

1. Let us consider the inner product space $(\mathbb{F}^n, \langle\cdot,\cdot\rangle_2)$. We represent an element $\boldsymbol{\alpha} \in (\mathbb{F}^n)^*$ by a $1 \times n$ matrix:
$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}(1) & \cdots & \boldsymbol{\alpha}(n) \end{bmatrix}.$$
The vector $\boldsymbol{v}_{\boldsymbol{\alpha}} \in \mathbb{F}^n$ corresponding to $\boldsymbol{\alpha}$ must then satisfy
$$\begin{aligned} & \boldsymbol{\alpha}(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{v}_{\boldsymbol{\alpha}} \rangle, \qquad \boldsymbol{u} \in \mathbb{F}^n \\ \implies \quad & \sum_{j=1}^n \boldsymbol{\alpha}(j)\boldsymbol{u}(j) = \sum_{j=1}^n \boldsymbol{u}(j)\overline{\boldsymbol{v}_{\boldsymbol{\alpha}}(j)}, \qquad \boldsymbol{u} \in \mathbb{F}^n \\ \implies \quad & \boldsymbol{v}_{\boldsymbol{\alpha}}(j) = \overline{\boldsymbol{\alpha}(j)}, \qquad j \in \{1, \dots, n\}. \end{aligned}$$
2. Next we consider the Hilbert space $(\ell^2(\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ and let $\alpha \in \ell^2(\mathbb{F})^*$. Then Corollary 7.2.5 ensures that there exists $\boldsymbol{v}_\alpha \in \ell^2(\mathbb{F})$ such that
$$\alpha(\boldsymbol{u}) = \sum_{j=1}^\infty \boldsymbol{u}(j)\overline{\boldsymbol{v}_\alpha(j)}$$
for every $\boldsymbol{u} \in \ell^2(\mathbb{F})$. From this expression we easily see that $\boldsymbol{v}_\alpha(j) = \overline{\alpha(\boldsymbol{e}_j)}$, $j \in \mathbb{Z}_{>0}$, where $\{\boldsymbol{e}_j\}_{j \in \mathbb{Z}_{>0}}$ is the standard basis for $\mathbb{F}_0^\infty$.
3. Finally, we consider the Hilbert space $(\mathsf{L}^2([a,b];\mathbb{F}), \langle\cdot,\cdot\rangle_2)$. If $\alpha \in \mathsf{L}^2([a,b];\mathbb{F})^*$ then Corollary 7.2.5 ensures that there exists $f_\alpha \in \mathsf{L}^2([a,b];\mathbb{F})$ such that
$$\alpha(g) = \int_a^b g(x) f_\alpha(x)\,\mathrm{d}x$$
for every $g \in \mathsf{L}^2([a,b];\mathbb{F})$. To extract a more explicit characterisation of f_α is possible once one has on hand the notion of a maximal orthonormal family. We refer to Exercise 7.3.8 for a working out of this characterisation. •

7.2.2 Particular aspects of continuity for inner product spaces

To get started we give a few constructions concerning linear maps between inner product spaces that are specific to the inner product structure. We begin with the notion of the adjoint of a continuous linear map.

7.2.7 Definition

7.2.8 Remark (Self-adjointness in Sturm–Liouville[2] theory) One of the important areas of application of inner product spaces is in so-called "Sturm–Liouville theory,"

[2] Friedrich Otto Rudolf Sturm (1841–1919) was a German mathematician whose contributions were mainly in the area of geometry. Joseph Liouville (1809–1882) was a French mathematician who made contributions to many areas of mathematics and its applications. These areas include mathematical physics, differential equations, number theory, and analysis.

which deals with a certain sort of ordinary differential equation. In this subject one is interested in linear maps that are self-adjoint. The sort of maps that arise in Sturm–Liouville theory are *not* of the sort coming from the preceding definition. There are many reasons why this is so, and we refer the reader to ***missing stuff*** for details. We mention this here because in reading some elementary treatments of Sturm–Liouville theory one might be led to believe that the theory has to do with the more or less simple situation of Definition 7.2.7. •

7.2.3 Notes

The Riesz Representation Theorem is frequently attributed to **FR:07c**, **FR:09** and also to **MF:07**.

Section 7.3

Orthonormal bases in Hilbert spaces

One of the features distinguishing Hilbert spaces from their more general Banach space brethren is that Hilbert spaces always possess a Schauder basis. In the theory of Hilbert space these bases go by various names, including maximal orthonormal set or complete orthonormal families; we use the former convention. The idea that every vector in a Hilbert space can be written as a (possibly infinite) sum of distinguished basis vectors is an important one, and plays an important rôle in the theory of, for example, Fourier series; see Chapter 12. Our presentation in this section begins with the finite-dimensional case in order to build some important intuition. We then progress to countable then general bases.

Do I need to read this section? This chapter, at least that part dealing with finite and countable maximal orthonormal sets, is important in our study of Fourier series in Chapter 12. Moreover, understanding the "geometry" of Hilbert spaces will be facilitated by understanding the notion of a maximal orthonormal set. •

7.3.1 General definitions and results

Before we proceed with our incremental treatment of orthonormal bases, let us give the definitions that apply to all inner product spaces.

7.3.1 Definition (Orthonormal set) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space.

(i) An ***orthogonal set*** is a collection $\{e_i\}_{i\in I}$ of nonzero vectors in V such that $\langle e_{i_1}, e_{i_2}\rangle = 0$ for all distinct $i_1, i_2 \in I$.

(ii) An ***orthonormal set*** is an orthogonal set $\{e_i\}_{i\in I}$ such that $\|e_i\| = 1$ for all $i \in I$. •

Sometimes we will talk about orthonormal and orthogonal families rather than sets. In this case we shall use the notation $(e_i)_{i\in I}$. The idea is the same, however.

Let us first indicate a useful construction for constructing orthonormal sets from linearly independent sets.

7.3.2 Theorem (Gram–Schmidt[3] orthonormalisation) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space, and let J be either the set $\{1, \dots, \mathrm{n}\}$ for some $\mathrm{n} \in \mathbb{Z}_{>0}$ or the set $\mathbb{Z}_{>0}$. For a family $(\mathrm{v_j})_{\mathrm{j}\in\mathrm{J}}$ of nonzero vectors in V define a family $(\mathrm{u'_j})_{\mathrm{j}\in\mathrm{J}}$ in V recursively by*

[3] Jorgen Pedersen Gram (1850–1916) was a Danish mathematician whose principal employer was the Hafnia Insurance Company. Much of Gram's mathematical work was devoted to using mathematical and statistical methods in forestry management. Despite being somewhat outside the main circle of activity in mathematics, Gram made real contributions to algebra, number theory, probability theory, and numerical analysis. Erhard Schmidt (1876–1959) was born in what is now Estonia. His principal mathematical contributions were to the areas of integral equations and functional analysis.

$\mathrm{u}'_1 = \mathrm{v}_1$ *and*

$$\mathrm{u}'_\mathrm{j} = \mathrm{v}_\mathrm{j} - \sum_{\mathrm{k}=1}^{\mathrm{j}-1} \frac{\langle \mathrm{v}_\mathrm{j}, \mathrm{u}'_\mathrm{k}\rangle}{\|\mathrm{u}'_\mathrm{k}\|^2}\mathrm{u}'_\mathrm{k}, \qquad \mathrm{j} \in \mathrm{J} \setminus \{1\}.$$

If the family $(\mathrm{v}_\mathrm{j})_{\mathrm{j}\in\mathrm{J}}$ *is linearly independent then the family* $(\mathrm{u}'_\mathrm{j})_{\mathrm{j}\in\mathrm{J}}$ *is orthogonal. Moreover, if we additionally define* $\mathrm{u}_\mathrm{j} = \frac{\mathrm{u}'_\mathrm{j}}{\|\mathrm{u}'_\mathrm{j}\|}$, $\mathrm{j} \in \mathrm{J}$, *then* $(\mathrm{u}_\mathrm{j})_{\mathrm{j}\in\mathrm{J}}$ *is orthonormal.*

Proof Let us prove that for any $m \in J$ the set $\{u'_1, \dots, u'_m\}$ is orthogonal. We prove this by induction on m. The claim is clearly true for $m = 1$. Suppose that the claim is true for $m = r$ so that $\{u'_1, \dots, u'_r\}$ is orthogonal. If $J = \{1, \dots, n\}$ and if $r = n$ then the claim is established. Otherwise we can carry on to show that $\{u'_1, \dots, u'_{r+1}\}$ is orthogonal as follows. For any $j \in \{1, \dots, r\}$,

$$\langle u'_{r+1}, u'_j\rangle = \Big\langle v_{r+1} - \sum_{k=1}^{r} \frac{\langle v_{r+1}, u'_k\rangle}{\|u'_k\|^2}u'_k, u'_j\Big\rangle = \langle v_{r+1}, u'_j\rangle - \langle v_{r+1}, u'_j\rangle = 0.$$

Thus u'_{r+1} is orthogonal to the set $\{u'_1, \dots, u'_r\}$. We claim that u'_{r+1} is nonzero. Indeed, by Exercise 7.3.1 we know that $\{u'_1, \dots, u'_r\}$ is linearly independent. Therefore,

$$\mathrm{span}_{\mathbb{F}}(v_1, \dots, v_r) = \mathrm{span}_{\mathbb{F}}(u'_1, \dots, u'_r).$$

Therefore, we have

$$u'_{r+1} = v_{r+1} + c_1 v_1 + \dots + c_r v_r$$

for $c_1, \dots, c_r \in \mathbb{F}$. If $u'_{r+1} = 0_\mathsf{V}$ then linear independence of $\{v_1, \dots, v_{r+1}\}$ gives $c_1 = \dots = c_r = 0$ and $1 = 0$. This last assertion is absurd, and so we must have $u'_{r+1} \neq 0_\mathsf{V}$. This shows that $\{u'_1, \dots, u'_{r+1}\}$ is indeed orthogonal.

Next we claim that orthogonality of $\{u'_1, \dots, u'_m\}$ for any $m \in J$ suffices to establish orthogonality of $(u'_j)_{j\in J}$. If J is finite this is obvious, so we consider the case where $J = \mathbb{Z}_{>0}$. In this case the family $(u'_j)_{j\in\mathbb{Z}_{>0}}$ could not be orthonormal in two ways.

1. One of the vectors u'_j, $j \in \mathbb{Z}_{>0}$, could be nonzero. This cannot happen, however, since for any $j \in \mathbb{Z}_{>0}$ the set $\{u_1, \dots, u_j\}$ is orthogonal.
2. For distinct $j_1, j_2 \in \mathbb{Z}_{>0}$ it could hold that $\langle u_{j_1}, u_{j_2}\rangle \neq 0$. This cannot happen, however, since for any distinct $j_1, j_2 \in \mathbb{Z}_{>0}$ the set $\{u_1, \dots, u_m\}$ is orthogonal for $m > \max\{j_1, j_2\}$.

The last assertion of the theorem is obvious. ■

7.3.3 Notation (Orthogonal sets) Generally we will use the notion of orthonormal set and not of an orthogonal set. However, in practice it is sometimes convenient to be able to talk about orthogonal sets as the objects which naturally present themselves are orthogonal, but not orthonormal. Note, however, that the two notions differ only in the trivial (but sometimes annoying) manner of nonzero constants. •

The following properties of orthonormal sets will be important to us in this section, and indeed in the study of inner product spaces in general.

7.3.4 Definition (Maximal, total, and basic orthonormal sets) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle \cdot, \cdot \rangle)$ be an $\mathbb{F}$-inner product space, and let $\{e_i\}_{i \in I}$ be an orthogonal (resp. orthonormal) set.

(i) The orthogonal (resp. orthonormal) set $\{e_i\}_{i \in I}$ is ***maximal*** if, for any orthogonal (resp. orthonormal) set $\{f_j\}_{j \in J}$ such that $\{e_i\}_{i \in I} \subseteq \{f_j\}_{j \in J}$, $\{f_j\}_{j \in J} \subseteq \{e_i\}_{i \in I}$.

(ii) The orthogonal (resp. orthonormal) set $\{e_i\}_{i \in I}$ is ***total*** if $\mathrm{cl}(\mathrm{span}_{\mathbb{F}}(\{e_i\}_{i \in I})) = \mathsf{V}$.

(iii) An orthogonal (resp. orthonormal) set $\{e_i\}_{i \in I}$ is ***basic*** if, for any $v \in \mathsf{V}$, there exist constants $c_i \in \mathbb{F}$, $i \in I$, for which the series

$$\sum_{i \in I} c_i e_i$$

converges to v in the sense of Definition 6.4.16. •

For convenience, let us recall here the definition of convergence used in the above definition for basic orthonormal sets. Convergence of the series

$$\sum_{i \in I} c_i e_i \tag{7.8}$$

to v means that, for every $\epsilon \in \mathbb{R}_{>0}$, there exists a finite set $J \subseteq I$ such that

$$\Big\| \sum_{j \in J} c_j e_j - v \Big\| < \epsilon.$$

By Proposition 6.4.18 it follows that a convergent sum of the form (7.8) is such that only countable many of the coefficients c_i, $i \in I$, are nonzero. Moreover, by Theorem 6.4.20, if the index set I is countable, say $I = \mathbb{Z}_{>0}$, then a sum

$$\sum_{j \in \mathbb{Z}_{>0}} c_j e_j$$

converges to v in the sense of Definition 6.4.16 if and only if it converges unconditionally to v. In particular, if this series converges to v in the sense of Definition 6.4.16 then it converges in the usual sense. It is usually the case that one deals with countable orthonormal sets.

Before we begin to explore properties of orthonormal sets of various flavours, let us give a few useful general results. First let us give the character of coefficients in any convergent series of orthonormal vectors.

7.3.5 Proposition (Coefficients in a convergent series of orthonormal vectors) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *let* $(\mathsf{V}, \langle \cdot, \cdot \rangle)$ *be an* $\mathbb{F}$*-inner product space, and let* $\{\mathrm{e_i}\}_{\mathrm{i \in I}}$ *be an orthonormal set. If the series*

$$\sum_{\mathrm{i \in I}} \mathrm{c_i e_i}$$

converges to $\mathrm{v} \in \mathsf{V}$ *then the coefficients must satisfy* $\mathrm{c_i} = \langle \mathrm{v}, \mathrm{e_i} \rangle$, $\mathrm{i} \in \mathrm{I}$.

Proof If I is finite then this is Exercise 7.3.4. Let us suppose, therefore, that I is infinite. Since the series converges, by Proposition 6.4.18 it follows that there exists an injection $\phi\colon \mathbb{Z}_{>0} \to I$ such that $c_i = 0$ for $i \notin \operatorname{image}(\phi)$ and such that

$$v = \sum_{j=1}^{\infty} c_{\phi(j)} e_{\phi(j)}.$$

Then, using Proposition 7.2.1 and Theorem 6.5.2, we deduce that for $j_0 \in \mathbb{Z}_{>0}$ we have

$$\langle v, e_{\phi(j_0)} \rangle = \Big\langle \sum_{j=1}^{\infty} c_j e_{\phi(j)}, e_{\phi(j_0)} \Big\rangle = \sum_{j=1}^{\infty} c_j \langle e_{\phi(j)}, e_{\phi(j_0)} \rangle = c_{\phi(j_0)},$$

giving $c_i = \langle v, e_i \rangle$ for $i \in \operatorname{image}(\phi)$. For $i \notin \operatorname{image}(\phi)$ a similar computation gives $\langle v, e_i \rangle = 0$¡ and so gives the result. ■

The following result is also useful.

7.3.6 Theorem (Bessel's[4] inequality) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is an $\mathbb{F}$-inner product space, if $\{\mathrm{e_i}\}_{\mathrm{i\in I}}$ is an orthonormal set, and if $\mathrm{v} \in \mathsf{V}$, then*

$$\sum_{\mathrm{i\in I}} |\langle \mathrm{v}, \mathrm{e_i} \rangle|^2 \le \|\mathrm{v}\|^2;$$

in particular, the sum on the left converges.

Proof By Exercise 7.3.5 we have

$$\sum_{j=1}^{n} |\langle v, e_{i_j} \rangle|^2 \le \|v\|^2$$

for every finite subset $\{i_1, \dots, i_n\} \subseteq I$. If I is finite this immediately gives the result. Let us consider the case where I is not finite. We claim that in this case $\langle v, e_i \rangle = 0$ for all but countably many $i \in I$. To see this, define

$$I_0 = \{i \in I \mid |\langle v, e_i \rangle| > 0\}$$

and suppose that I_0 is not countable. For $k \in \mathbb{Z}_{>0}$ define

$$I_k = \{i \in I \mid |\langle v, e_i \rangle|^2 \ge \tfrac{1}{k}\}.$$

Note that $I_0 = \cup_{k \in \mathbb{Z}_{>0}} I_k$, implying by Proposition **??** that for at least one $k \in \mathbb{Z}_{>0}$ the set I_k must be infinite (uncountable, actually, although this is not necessary). Let $N \in \mathbb{Z}_{>0}$ be such that $N > k\|v\|^2$. Then, for any finite subset $\{i_1, \dots, i_N\} \subseteq I_k$ we have

$$\sum_{j=1}^{N} |\langle v, e_{i_j} \rangle|^2 \ge \sum_{j=1}^{N} \frac{1}{k} = \frac{N}{k} > \|v\|^2,$$

which gives a contradiction. Thus I_0 must indeed be countable.

[4] Friedrich Wilhelm Bessel (1784–1846) was born in what is now Germany and made mathematical contributions to analysis. His primary scientific activities were directed towards astronomy.

Thus we have an injection $\phi\colon \mathbb{Z}_{>0} \to I$ such that $\langle v, e_i \rangle = 0$ for $i \notin \operatorname{image}(\phi)$ and such that

$$\sum_{j=1}^{\infty} |\langle v, e_{\phi(j)} \rangle|^2 = \lim_{n\to\infty} \sum_{j=1}^{n} |\langle v, e_{\phi(j)} \rangle|^2 \leq \|v\|^2.$$

Thus we have

$$\sum_{i\in I} |\langle v, e_i \rangle|^2 \leq \|v\|^2$$

for every index set I. Since this is a sum of positive terms, the series

$$\sum_{i\in I} |\langle v, e_i \rangle|^2$$

converges for arbitrary index sets I. ■

Bessel's inequality makes the following definition reasonable.

7.3.7 Definition (Orthonormal expansion) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space, and let $\{e_i\}_{i\in I}$ be an orthonormal set. The ***orthonormal expansion*** of $v \in \mathsf{V}$ with respect to $\{e_i\}_{i\in I}$ is the series

$$\sum_{i\in I} \langle v, e_i \rangle e_i,$$

disregarding convergence. •

Let us give some examples of orthonormal sets.

7.3.8 Examples (Orthonormal sets)

1. In $\mathbb{F}^n$ with the standard inner product, one can check that the standard basis,

$$\boldsymbol{e}_1 = (1, 0, \dots, 0),\ \boldsymbol{e}_2 = (0, 1, \dots, 0),\ \dots,\ \boldsymbol{e}_n = (0, 0, \dots, 1),$$

is orthonormal. That is to say, the set $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is orthonormal. Moreover, the set $\{\lambda_1\boldsymbol{e}_1, \dots, \lambda_n\boldsymbol{e}_n\}$ is orthogonal for any collection of constants $\lambda_1, \dots, \lambda_n \in \mathbb{F}\setminus\{0\}$. Orthonormality of this set occurs precisely when $\lambda_j = 1$, $j \in \{1, \dots, n\}$.

It is easy to see that $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is a maximal orthonormal set. Indeed, let us consider an orthonormal set $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n, \boldsymbol{e}_{n+1}, \dots, \boldsymbol{e}_k\}$ containing $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. We claim that $k = n$. Suppose otherwise. Since $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is a basis for $\mathbb{F}^n$ it follows that for each $a \in \{n+1, \dots, k\}$,

$$\boldsymbol{e}_a = c_{a1}\boldsymbol{e}_1 + \dots + c_{an}\boldsymbol{e}_n$$

for some constants $c_{a1}, \dots, c_{an}$. Since $\langle \boldsymbol{e}_a, \boldsymbol{e}_j \rangle = 0$ it follows that $c_{aj} = 0$ for $a \in \{n+1, \dots, k\}$ and $j \in \{1, \dots, n\}$. Thus $\boldsymbol{e}_{n+1} = \dots = \boldsymbol{e}_k = \boldsymbol{0}$, contradicting the orthonormality of $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$. Thus $k = n$.

Moreover, since $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is a basis for $\mathbb{F}^n$ it follows that $\operatorname{span}_{\mathbb{F}}(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) = \mathbb{F}^n$, and so the orthonormal set is total and basic.

2. Next let us consider the inner product space $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$. We note that the standard basis $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$, which we recall is defined by

$$\boldsymbol{e}_j(k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k, \end{cases}$$

is orthonormal; this is straightforward to verify. Moreover, the set $(\lambda_j\boldsymbol{e}_j)_{j\in\mathbb{Z}_{>0}}$ is orthogonal for every collection of constants $\lambda_j \in \mathbb{F}\setminus\{0\}$, $j \in \mathbb{Z}_{>0}$, and is orthonormal if and only if $\lambda_j = 1$, $j \in \mathbb{Z}_{>0}$.

We leave it to the reader to show in Exercise 7.3.2 to show that $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ is a maximal orthonormal family.

Moreover, since $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ is a basis for $\mathbb{F}_0^\infty$ it follows that $\mathrm{span}_{\mathbb{F}}(\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}) = \mathbb{F}_0^\infty$, and so the orthonormal set is total and basic.

3. The preceding examples might make one believe that the notions of maximal, total, and basic orthonormal sets are equivalent for general inner product spaces. They are not. Let us give an example to illustrate this. We consider the Hilbert space $(\ell^2(\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ with $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ the orthonormal set from the preceding example, i.e., the standard (Hamel) basis for $\mathbb{F}_0^\infty \subseteq \ell^2(\mathbb{F})$. We then take the subspace

$$\mathsf{U} = \mathrm{span}_{\mathbb{F}}\Big(\sum_{j=1}^{\infty}\frac{\boldsymbol{e}_j}{j}, \boldsymbol{e}_2, \boldsymbol{e}_3, \dots\Big)$$

and consider $(\mathsf{U}, \langle\cdot,\cdot\rangle_2)$ as an inner product space. We claim that $\mathscr{B} = \{\boldsymbol{e}_2, \boldsymbol{e}_3, \dots\}$ is a maximal orthonormal set in U that is neither total nor basic.

To show that it is maximal, suppose that $\boldsymbol{u} \in \mathsf{U}$ is orthogonal to $\mathscr{B}$. Since $\boldsymbol{u} \in \mathsf{U}$ we can write

$$\boldsymbol{u} = c_1\Big(\sum_{j=1}^{\infty}\frac{\boldsymbol{e}_j}{j}\Big) + c_2\boldsymbol{e}_2 + \dots + c_k\boldsymbol{e}_k$$

for some $k \in \mathbb{Z}_{>0}$ and for $c_1, \dots, c_k \in \mathbb{F}$. Since

$$\Big\langle \boldsymbol{u}, \sum_{j=1}^{\infty}\frac{\boldsymbol{e}_j}{j}\Big\rangle = 0, \ \langle \boldsymbol{u}, \boldsymbol{e}_j\rangle = 0, \qquad j \in \{2, 3, \dots\},$$

it follows that $c_j = 0$, $j \in \{1, \dots, k\}$, and so $\boldsymbol{u} = \boldsymbol{0}_{\mathbb{F}_0^\infty}$. Thus there can be no orthonormal subset of U containing $\mathscr{B}$.

That $\mathscr{B}$ is not basic is plain since $\sum_{j=1}^{\infty}\frac{e_j}{j}$ is in U but is not a sum of the form $\sum_{j=2}^{\infty} c_j\boldsymbol{e}_j$ (this follows from Proposition 7.3.5).

That $\mathscr{B}$ is not total follows since the subspace $\mathrm{span}_{\mathbb{F}}(\boldsymbol{e}_2, \boldsymbol{e}_3, \dots)$ is a closed subspace containing $\mathscr{B}$ but is a strict subspace of V. •

The preceding examples illustrate that the notions of maximal, total, and basic need not be equivalent for an orthonormal set. Let us explore the relationships between these concepts in a general setting.

7.3.9 Theorem (Relationship between maximal, total, and basic orthonormal sets) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space, and let $\mathscr{B} = \{\mathrm{e}_\mathrm{i}\}_{\mathrm{i}\in\mathrm{I}}$ be an orthonormal set. The following four statements are equivalent:*

(i) $\mathscr{B}$ is basic;

(ii) $\mathscr{B}$ is total;

(iii) for every $\mathrm{v} \in \mathsf{V}$ the equality

$$\|\mathrm{v}\|^2 = \sum_{\mathrm{i}\in\mathrm{I}} |\langle \mathrm{v}, \mathrm{e}_\mathrm{i}\rangle|^2$$

*holds, where convergence of the sum on the right is interpreted as in Definition 2.4.31 (**Parseval's equality**);*

(iv) for all $\mathrm{u}, \mathrm{v} \in \mathsf{V}$ we have

$$\langle \mathrm{u}, \mathrm{v}\rangle = \sum_{\mathrm{i}\in\mathrm{I}} \langle \mathrm{u}, \mathrm{e}_\mathrm{i}\rangle \overline{\langle \mathrm{v}, \mathrm{e}_\mathrm{i}\rangle},$$

where convergence of the sum on the right is interpreted as in Definition 2.4.31.

Also, the following two statements are equivalent:

(v) $\mathscr{B}^\perp = \{0_\mathsf{V}\}$;

(vi) $\mathscr{B}$ is maximal.

Finally, if V is a Hilbert space, the first four equivalent statements are equivalent to the last two equivalent statements.

Proof (i) $\Longrightarrow$ (ii) Let $\mathscr{B} = \{e_i\}_{i\in I}$ be basic and let $v \in \mathsf{V}$. We can then write

$$v = \sum_{i\in I} c_i e_i$$

for some coefficients $c_i \in \mathbb{F}$, $i \in I$. If I is finite this immediately implies that $v \in \mathrm{cl}(\mathrm{span}_\mathbb{F}(\mathscr{B}))$. If I is not finite, by Proposition 6.4.18 and Theorem 6.4.20 there exists an injection $\phi\colon \mathbb{Z}_{>0} \to I$ such that $c_i = 0$ for $i \notin \mathrm{image}(\phi)$ and such that

$$v = \sum_{j=1}^\infty c_j e_{\phi(j)}.$$

If we define

$$v_k = \sum_{j=1}^k c_j e_j$$

then the sequence $(v_k)_{k\in\mathbb{Z}_{>0}}$ converges to v. Thus $v \in \mathrm{cl}(\mathrm{span}_\mathbb{F}(\mathscr{B}))$ and so $\mathscr{B}$ is total.

(ii) $\Longrightarrow$ (iii) Let $v \in \mathsf{V}$. Since $\mathscr{B}$ is total there exists a sequence $(v_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathrm{span}_\mathbb{F}(\mathscr{B})$ such that $v = \lim_{j\to\infty} v_j$. For each $j \in \mathbb{Z}_{>0}$ write

$$v_j = c_{j1}e_{i_{j1}} + \cdots + c_{jk_j}e_{i_{jk_j}}$$

for $k_j \in \mathbb{Z}_{>0}$, coefficients $c_{j1}, \dots, c_{jk_j} \in \mathbb{F}$, and distinct $i_{j1}, \dots, i_{jk_j} \in I$. By Exercise 7.3.4 it follows that $c_{jl} = \langle v_j, e_{i_{jl}}\rangle$ for each $j \in \mathbb{Z}_{>0}$, $l \in \{1, \dots, k_j\}$. Note that the set $\cup_{j\in\mathbb{Z}_{>0}}\{i_{j1},\dots,i_{jk_j}\}$ is countable by Proposition **??**. This means that there exists a countable set $K \subseteq I$ such that

$$v_j = \sum_{k\in K}\langle v_j, e_k\rangle e_k$$

for each $j \in \mathbb{Z}_{>0}$, with the sum being finite. We claim that $\langle v, e_i\rangle = 0$ for $i \notin K$. Indeed, for $i \in I$,

$$\langle v, e_i\rangle = \lim_{j\to\infty}\langle v_j, e_i\rangle = \lim_{j\to\infty}\sum_{k\in K}\langle v_j, e_k\rangle\langle e_k, e_i\rangle = 0,$$

using continuity of the inner product and Theorem 6.5.2.

We now have

$$\begin{aligned}\|v_j\|^2 &= \Big\langle \sum_{k\in K}\langle v_j, e_k\rangle e_k, \sum_{k'\in K}\langle v_j, e_{k'}\rangle e_{k'}\Big\rangle\\ &= \sum_{k\in K}\sum_{k'\in K}\langle v_j, e_k\rangle\overline{\langle v_j, e_{k'}\rangle}\langle e_k, e_{k'}\rangle\\ &= \sum_{k\in K}|\langle v_j, e_k\rangle|^2,\end{aligned}$$

using the fact that the inner product commutes with finite sums. Now, using continuity of the norm and inner product, along with Theorem 6.5.2, gives

$$\|v\|^2 = \lim_{j\to\infty}\|v_j\|^2 = \lim_{j\to\infty}\sum_{k\in K}|\langle v_j, e_k\rangle|^2 = \sum_{k\in K}|\langle v, e_k\rangle|^2 = \sum_{i\in I}|\langle v, e_i\rangle|^2,$$

as desired.

(iii) $\Longrightarrow$ (iv) For $u, v \in \mathsf{V}$ we have

$$\begin{aligned} & \|u+v\|^2 = \sum_{i\in I}|\langle u+v, e_i\rangle|^2\\ \Longrightarrow\quad & \|u\|^2 + \|v\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle}\\ & = \sum_{i\in I}|\langle u, e_i\rangle|^2 + \sum_{i\in I}|\langle v, e_i\rangle|^2 + \sum_{i\in I}(\langle u, e_i\rangle\overline{\langle v, e_i\rangle} + \overline{\langle u, e_i\rangle\overline{\langle v, e_i\rangle}})\\ \Longrightarrow\quad & \operatorname{Re}(\langle u, v\rangle) = \sum_{i\in I}\operatorname{Re}(\langle u, e_i\rangle\overline{\langle v, e_i\rangle}).\end{aligned}$$

If $\mathbb{F} = \mathbb{R}$ this establishes the result. If $\mathbb{F} = \mathbb{C}$, a similar computation using the equality

$$\|u + iv\|^2 = \sum_{i\in I}|\langle u + iv, e_i\rangle|^2$$

gives

$$\operatorname{Im}(\langle u, v\rangle) = \sum_{i\in I}\operatorname{Im}(\langle u, e_i\rangle\overline{\langle v, e_i\rangle}).$$

(iv) $\Longrightarrow$ (i) Since part (iv) obviously implies part (iii), we shall prove that (iii) implies (i). Thus we have

$$\|v\|^2 = \sum_{i\in I}|\langle v, e_i\rangle|^2$$

for every $v \in \mathsf{V}$. By Proposition 2.4.33, for $v \in \mathsf{V}$, it follows that there exists a bijection $\phi\colon \mathbb{Z}_{>0} \to I$ such that $\langle v, e_i\rangle = 0$ for $i \notin \operatorname{image}(\phi)$ and such that

$$\|v\|^2 = \sum_{j=1}^{\infty} |\langle v, e_{\phi(j)}\rangle|^2.$$

For $k \in \mathbb{Z}_{>0}$ let us define

$$v_k = \sum_{j=1}^{k} \langle v, e_{\phi(j)}\rangle e_{\phi(j)}.$$

Note that

$$\begin{aligned}\langle v - v_k, v_k\rangle &= \Big\langle v - \sum_{j=1}^{k}\langle v, e_{\phi(j)}\rangle e_{\phi(j)}, \sum_{l=1}^{k}\langle v, e_{\phi(l)}\rangle e_{\phi(l)}\Big\rangle\\ &= \Big\langle v, \sum_{l=1}^{k}\langle v, e_{\phi(l)}\rangle e_{\phi(l)}\Big\rangle - \Big\langle \sum_{j=1}^{k}\langle v, e_{\phi(j)}\rangle e_{\phi(j)}, \sum_{l=1}^{k}\langle v, e_{\phi(l)}\rangle e_{\phi(l)}\Big\rangle\\ &= \sum_{l=1}^{k}|\langle v, e_{\phi(l)}\rangle|^2 - \sum_{j=1}^{k}|\langle v, e_{\phi(j)}\rangle|^2 = 0\end{aligned}$$

for every $k \in \mathbb{Z}_{>0}$. By the Pythagorean equality,

$$\|v\|^2 = \|v - v_k + v_k\|^2 = \|v - v_k\|^2 + \|v_k\|^2 \quad\Longrightarrow\quad \|v - v_k\|^2 = \|v\|^2 - \|v_k\|^2.$$

By assumption,

$$\lim_{k\to\infty} \|v_k\|^2 = \|v\|^2$$

and so

$$\lim_{k\to\infty} \|v - v_k\| = 0,$$

implying that

$$v = \sum_{i\in I} \langle v, e_i\rangle e_i,$$

and so in particular implying that $\mathscr{B}$ is basic.

(v) $\Longrightarrow$ (vi) Suppose that $\mathscr{B}$ is not maximal. Then there exists an orthonormal set $\mathscr{B}'$ such that $\mathscr{B} \subset \mathscr{B}'$. Let $v \in \mathscr{B}' \setminus \mathscr{B}$. Then, clearly, $v \in \mathscr{B}^\perp$ and $v \neq 0_{\mathsf{V}}$. Thus $\mathscr{B}^\perp \neq \{0_{\mathsf{V}}\}$.

(vi) $\Longrightarrow$ (v) Suppose that $\mathscr{B}^\perp \neq \{0_{\mathsf{V}}\}$ and let $v \in \mathscr{B}^\perp$ have unit length. Then the set $\mathscr{B} \cup \{v\}$ is an orthonormal set that strictly contains $\mathscr{B}$. Thus $\mathscr{B}$ is not maximal.

(ii) $\Longrightarrow$ (v) By Proposition 7.1.13(iv) we have $\mathscr{B}^\perp = \operatorname{cl}(\operatorname{span}_{\mathbb{F}}(\mathscr{B}))^\perp$. From this fact, if $\mathscr{B}$ is total it immediately follows that $\mathscr{B}^\perp = \{0_{\mathsf{V}}\}$.

(vi) $\Longrightarrow$ (i) (assuming V is a Hilbert space) Let $v \in \mathsf{V}$. Bessel's inequality gives

$$\sum_{i\in I} |\langle v, e_i\rangle|^2 \leq \|v\|^2,$$

and this implies that the series on the right converges and so is Cauchy. Let $\epsilon \in \mathbb{R}_{>0}$ and let $J \subseteq I$ be a finite set for which

$$\sum_{j \in J'} |\langle v, e_j \rangle|^2 < \epsilon$$

for every finite subset $J' \subseteq I$ such that $J \cap J' = \emptyset$ (see Definition 6.4.16). A direct computation using properties of inner products then gives

$$\Big\| \sum_{j \in J'} \langle v, e_j \rangle e_j \Big\|^2 = \sum_{j \in J'} |\langle v, e_j \rangle|^2 < \epsilon,$$

which shows that the series

$$\sum_{i \in I} \langle v, e_i \rangle e_i$$

is Cauchy. By Theorem 6.4.17 this series converges, implying that $\mathscr{B}$ is basic. ■

The following result records the fact that completeness is essential if all six of the statements in the preceding theorem are to be equivalent.

7.3.10 Theorem (Maximal orthonormal sets are not generally Hilbert bases for non-Hilbert spaces) *For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and for a $\mathbb{F}$-inner product space $(\mathsf{V}, \langle \cdot, \cdot \rangle)$, the following statements are equivalent:*

(i) V is a Hilbert space;

(ii) every maximal orthonormal set is a Hilbert basis.

Proof The implication of part (ii) from part (i) follows from Theorem 7.3.9, so we prove the converse implication. Let U be a proper closed subspace of V. By Theorem 7.1.23, to show that V is a Hilbert space it suffices to show that $\mathsf{U}^\perp \neq \{0_{\mathsf{V}}\}$. So suppose otherwise. Now let $\mathscr{B} = \{e_i\}_{i \in I}$ be a maximal orthonormal set in U and let $\mathscr{B}' = \mathscr{B} \cup \{f_j\}_{j \in J}$ be a maximal orthonormal set in V that extends that $\mathscr{B}$ (that such a set exists may be proved just as one proves Theorem 4.3.26). Let $j_0 \in J$. Since $f_{j_0} \neq 0_{\mathsf{V}}$ it follows that $f_{j_0} \notin \mathsf{U}^\perp$. Thus there exists $u \in \mathsf{U}$ such that $\langle u, f_{j_0} \rangle \neq 0$. By hypothesis, $\mathscr{B}'$ is a basic orthonormal set and so we may write

$$u = \sum_{i \in I} a_i e_i + \sum_{j \in J} b_j f_j$$

for some coefficients $a_i \in \mathbb{F}$, $i \in I$, $b_j \in \mathbb{F}$, $j \in J$. Then

$$\sum_{j \in J} b_j f_j = u - \sum_{i \in I} a_i e_i \in \mathsf{U}.$$

We also have

$$\sum_{j \in J} b_j f_j \in \mathscr{B}^\perp.$$

Since $\mathscr{B}$ is a maximal orthonormal set in U it follows that

$$\sum_{j \in J} b_j f_j = 0_{\mathsf{V}}$$

and so

$$\langle u, f_{j_0}\rangle = \Big\langle \sum_{i\in I} a_i e_i, f_{j_0}\Big\rangle = \sum_{i\in I} a_i \langle e_i, f_{j_0}\rangle = 0,$$

where we have used Proposition 7.2.1. This is a contradiction. Thus it must be the case that $\mathsf{U}^\perp \neq \{0_{\mathsf{V}}\}$. ■

Now that we have a clear understanding of the relationships between basic, total, and maximal orthonormal sets, let us introduce some useful terminology.

7.3.11 Definition (Hilbert basis) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. A ***Hilbert basis*** for an $\mathbb{F}$-inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is a basic (or, equivalently, total) orthonormal set in V. •

As we shall see, the notion of a Hilbert basis and a basis (sometimes also called a Hamel basis, cf. Remark 4.3.21) can be different in a potentially confusing way. In particular, we refer to ***missing stuff*** to clarify some aspects of the relationship between the two notions of basis.

We have already seen in Example 7.3.8–3 that not every inner product space possesses a Hilbert basis. This, however, is where the value of the notion of a maximal orthonormal set arises.

7.3.12 Theorem (Every inner product spaces possesses a maximal orthonormal set) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is a $\mathbb{F}$-inner product space, then there exists a maximal orthonormal set in V.*

Proof The proof goes very much like that for existence of a (Hamel) basis. Let $\mathscr{O}$ be the collection of orthonormal subsets of V. This set is nonempty since, if V is not the trivial vector space, $\{v\} \in \mathscr{O}$ for any vector v of unit length. Place a partial order $\preceq$ on $\mathscr{O}$ by asking that $S_1 \preceq S_2$ if $S_1 \subseteq S_2$. Let $\mathscr{S} \subseteq \mathscr{O}$ be a totally ordered subset. Note that $\cup_{S\in\mathscr{S}} S$ is an element of $\mathscr{O}$. Indeed, let $\{v_1, \dots, v_k\} \subseteq \cup_{S\in\mathscr{S}} S$. Then $v_j \in S_j$ for some $S_j \in \mathscr{S}$. Let $j_0 \in \{1, \dots, k\}$ be chosen such that S_{j_0} is the largest of the sets $S_1, \dots, S_k$ according to the partial order $\preceq$, this being possible since $\mathscr{S}$ is totally ordered. Then $\{v_1, \dots, v_k\} \subseteq S_{j_0}$ and so $\{v_1, \dots, v_k\}$ is orthonormal since S_{j_0} is orthonormal. It is also evident that $\cup_{S\in\mathscr{S}} S$ is an upper bound for $\mathscr{S}$. Thus every totally ordered subset of $\mathscr{O}$ possesses an upper bound, and so by Zorn's Lemma possesses a maximal element. Let $\mathscr{B}$ be such a maximal element. By construction $\mathscr{B}$ is orthonormal. We claim that it is also a maximal orthonormal set. Indeed, let $\mathscr{B}'$ be an orthonormal set such that $\mathscr{B} \subseteq \mathscr{B}'$. This immediately contradicts the fact that $\mathscr{B}$ is a maximal element of $\mathscr{O}$, and so we can conclude that $\mathscr{B}$ is a maximal orthonormal set. ■

For Hilbert spaces this leads to the following important result.

7.3.13 Corollary (Hilbert spaces possess a Hilbert basis) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is an $\mathbb{F}$-Hilbert space, then there exists a Hilbert basis for V.*

Proof By Theorem 7.3.12 V possesses a maximal orthonormal set. By Theorem 7.3.9 every maximal orthonormal set is a Hilbert basis. ■

Note that it is not necessary for an inner product space to be a Hilbert space in order that it possess a Hilbert basis, cf. Example 7.3.8–2.

Now we consider a few important special cases of inner product spaces with orthonormal bases. While many of the result we give in the next two sections are actually special cases of the results above, we give independent proofs that are not dependent on the notion of a sum with an arbitrary index set. The relieves some of the complication present in the general setup.

7.3.2 Finite orthonormal sets and finite Hilbert bases

In this section we essentially generalise Example 7.3.8–1 to arbitrary finite-dimensional inner product spaces. The starting point is the following result. Note that we independently prove the existence of a Hilbert basis in this case, although this actually follows from Theorem 7.3.12.

7.3.14 Theorem (Characterisation of existence of finite Hilbert bases) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is an inner product space of dimension* $\mathrm{n} \in \mathbb{Z}_{\geq 0}$*, then there exists a Hilbert basis for* V*. Moreover, every Hilbert basis for* V *is a basis and so has cardinality* n*.*

Proof If $\mathsf{V} = \{0_{\mathsf{V}}\}$ then there is nothing to prove, so let us suppose that $n \in \mathbb{Z}_{>0}$. By Theorem 4.3.22 V possesses a basis and by Theorem 4.3.25 the cardinality of any two bases are the same. Let $\{v_1, \dots, v_n\}$ be a basis for V and by Gram–Schmidt orthonormalisation construct an orthonormal set $\{u_1, \dots, u_n\}$. This set is linearly independent by Exercise 7.3.1 and so forms a basis for an n-dimensional subspace of V. By Proposition 4.3.19 this subspace must be V. That is to say, $\{u_1, \dots, u_n\}$ is a basis for V. We claim that this implies that $\{u_1, \dots, u_n\}$ is a *Hilbert* basis. Since finite-dimensional inner product spaces are Hilbert spaces, it suffices to show that $\mathscr{B}$ is maximal. To prove maximality, suppose that $\{u_1, \dots, u_n, u_{n+1}, \dots, u_k\}$ is an orthonormal set containing $\{u_1, \dots, u_n\}$. By Exercise 7.3.1 it follows that $\{u_1, \dots, u_k\}$ is linearly independent. By Lemma 1 from the proof of Theorem 4.3.25 it follows that $k = n$, so proving maximality. This gives the existence of a Hilbert basis.

For the last assertion of the theorem, suppose that we have a Hilbert basis $\{u_1, \dots, u_m\}$ for V. Since $\{u_1, \dots, u_m\}$ is linearly independent by Exercise 7.3.1 it follows that $m \leq n$ by Lemma 1 from the proof of Theorem 4.3.25. To see that $m = n$ suppose otherwise so that $n > m$. Then $\mathrm{span}_{\mathbb{F}}(u_1, \dots, u_m)$ is a subspace of V of dimension $m < n$. By Theorem 4.3.26 there exists $u_{m+1}, \dots, u_n \in \mathsf{V}$ such that $\{u_1, \dots, u_n\}$ is a basis for V. Applying the Gram–Schmidt orthonormalisation procedure gives a set $\{u'_1, \dots, u'_m, u'_{m+1}, \dots, u'_n\}$ where, by Exercise 7.3.3, $u'_j = u_j$ for $j \in \{1, \dots, m\}$. This contradicts the maximality of $\{u_1, \dots, u_m\}$ and so shows that we must have $m = n$. Thus every Hilbert basis is a linearly independent set of vectors having the same cardinality as the dimension of V, i.e., a basis. ■

A companion to the preceding result is the following more or less obvious fact.

7.3.15 Proposition (Necessary conditions for a finite Hilbert basis) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is an* $\mathbb{F}$*-inner product space having a finite Hilbert basis, then* V *is finite-dimensional.*

Proof Let $\{e_1, \dots, e_n\}$ be a finite Hilbert basis for V. We claim that $\dim(\mathsf{V}) = n$. Suppose otherwise. Then, by Theorem 4.3.26, there exists a basis $\mathscr{B}$ for V such that $\{e_1, \dots, e_n\} \subset \mathscr{B}$. Let $v \in \mathscr{B} \setminus \{e_1, \dots, e_n\}$. By applying Gram–Schmidt orthonormalisation procedure to $\{e_1, \dots, e_n, v\}$ we arrive at an orthonormal set $\{e_1, \dots, e_n, e_{n+1}\}$; by virtue of Exercise 7.3.3 the first n vectors remain unchanged. This, however, contradicts the maximality of $\{e_1, \dots, e_n\}$, and so we must have $\dim(\mathsf{V}) = n$. ■

Having established the existence of a Hilbert basis for a finite-dimensional inner product space, let us examine the set of all such bases. To motivate how one does this, recall from Section **??** that there is a 1–1 correspondence between bases and invertible matrices. That is to say, if one chooses a basis $\mathscr{B} = \{e_1, \dots, e_n\}$ for V, then any other basis $\mathscr{B}' = \{e'_1, \dots, e'_n\}$ is uniquely determined by the invertible change of basis matrix $\boldsymbol{P}^{\mathscr{B}'}_{\mathscr{B}} \in \mathrm{Mat}_{n\times n}(\mathbb{F})$ which is defined by its satisfying the equality

$$e_{j_0} = \sum_{j=1}^{n} \boldsymbol{P}^{\mathscr{B}'}_{\mathscr{B}}(j, j_0) e'_j$$

for each $j_0 \in \{1, \dots, n\}$. We wish to understand the character of the change of basis matrix in the case where $\mathscr{B}$ and $\mathscr{B}'$ are both Hilbert bases.

The following result tells the story. In the statement, $\langle\cdot,\cdot\rangle_2$ denotes the standard inner product on $\mathbb{F}^n$ and $\|\cdot\|_2$ denotes the corresponding norm. Also, for a matrix $\boldsymbol{A}$ we denote by $\bar{\boldsymbol{A}}$ the matrix obtained by applying $\bar{\cdot}$ to the entries of $\boldsymbol{A}$.

7.3.16 Theorem (Change of basis matrices for finite Hilbert bases) *For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, for an n-dimensional $\mathbb{F}$-inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$, for a Hilbert basis $\mathscr{B} = \{\mathrm{e}_1, \dots, \mathrm{e}_\mathrm{n}\}$ for V, and for $\mathbf{U} \in \mathrm{Mat}_{\mathrm{n}\times\mathrm{n}}(\mathbb{F})$ the following statements are equivalent:*

(i) there exists a Hilbert basis $\mathscr{B}' = \{\mathrm{e}'_1, \dots, \mathrm{e}'_\mathrm{n}\}$ for V such that $\mathbf{U} = \mathbf{P}^{\mathscr{B}'}_{\mathscr{B}}$;

(ii) $\|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ for all $\mathbf{x} \in \mathbb{F}^\mathrm{n}$;

(iii) $\langle \mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y}\rangle_2 = \langle \mathbf{x}, \mathbf{y}\rangle_2$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^\mathrm{n}$;

(iv) $\mathbf{U}\bar{\mathbf{U}}^\mathrm{T} = \bar{\mathbf{U}}^\mathrm{T}\mathbf{U} = \mathbf{I}_\mathrm{n}$;

(v) $\mathbf{U}$ is invertible and $\mathbf{U}^{-1} = \bar{\mathbf{U}}^\mathrm{T}$.

Proof (i) $\Longrightarrow$ (ii) By hypothesis we have

$$e_{j_0} = \sum_{j=1}^{n} \boldsymbol{U}(j, j_0) e'_j, \qquad j_0 \in \{1, \dots, n\},$$

so that, for every $j_1, j_2 \in \{1, \dots, n\}$,

$$\langle e_{j_1}, e_{j_2}\rangle = \Big\langle \sum_{k=1}^{n} \boldsymbol{U}(k, j_1) e'_k, \sum_{l=1}^{n} \boldsymbol{U}(l, j_2) e'_l \Big\rangle = \sum_{k=1}^{n} \boldsymbol{U}(k, j_1)\bar{\boldsymbol{U}}(k, j_2). \tag{7.9}$$

That is,

$$\sum_{k=1}^{n} \boldsymbol{U}(k, j_1)\bar{\boldsymbol{U}}(k, j_2) = \begin{cases} 1, & j_1 = j_2, \\ 0, & j_1 \neq j_2. \end{cases} \tag{7.10}$$

Now, for $\boldsymbol{x} \in \mathbb{F}^n$, a direct computation gives

$$\|\boldsymbol{U}\boldsymbol{x}\|_2^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \boldsymbol{U}(i, j)\bar{\boldsymbol{U}}(i, k)\boldsymbol{x}(j)\boldsymbol{x}(k)$$

which gives $\|\boldsymbol{U}\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2$ after using (7.10). This part of the result now follows by taking square roots.

(ii) $\implies$ (iii) We are assuming that $\|\boldsymbol{U}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$ which implies that

$$\|\boldsymbol{U}\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2 \quad\implies\quad \langle \boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{x}\rangle_2 = \langle \boldsymbol{x}, \boldsymbol{x}\rangle_2,$$

this holding for all $\boldsymbol{x} \in \mathbb{F}^n$. Thus, for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$,

$$\begin{aligned} & \langle \boldsymbol{U}(\boldsymbol{x}+\boldsymbol{y}), \boldsymbol{U}(\boldsymbol{x}+\boldsymbol{y})\rangle_2 = \langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle_2 \\ \implies\quad & \langle \boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{x}\rangle_2 + \langle \boldsymbol{U}\boldsymbol{y}, \boldsymbol{U}\boldsymbol{y}\rangle_2 + 2\,\mathrm{Re}(\langle \boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{y}\rangle_2) = \langle \boldsymbol{x}, \boldsymbol{x}\rangle_2 + \langle \boldsymbol{y}, \boldsymbol{y}\rangle_2 + 2\,\mathrm{Re}(\langle \boldsymbol{x}, \boldsymbol{y}\rangle_2) \\ \implies\quad & \mathrm{Re}(\langle \boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{y}\rangle_2) = \mathrm{Re}(\langle \boldsymbol{x}, \boldsymbol{y}\rangle_2). \end{aligned}$$

If $\mathbb{F} = \mathbb{R}$ then this gives this part of the result. If $\mathbb{F} = \mathbb{C}$, a computation entirely similar to the preceding one shows that

$$\langle \boldsymbol{U}(\boldsymbol{x}+\mathrm{i}\boldsymbol{y}), \boldsymbol{U}(\boldsymbol{x}+\mathrm{i}\boldsymbol{y})\rangle_2 = \langle \boldsymbol{x}+\mathrm{i}\boldsymbol{y}, \boldsymbol{x}+\mathrm{i}\boldsymbol{y}\rangle_2 \implies \mathrm{Im}(\langle \boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{y}\rangle_2) = \mathrm{Im}(\langle \boldsymbol{x}, \boldsymbol{y}\rangle_2),$$

which gives this part of the result.

(iii) $\implies$ (iv) Letting $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ be the standard basis for $\mathbb{F}^n$ we have

$$\langle \boldsymbol{U}\boldsymbol{e}_j, \boldsymbol{U}\boldsymbol{e}_k\rangle_2 = \langle \boldsymbol{e}_j, \boldsymbol{e}_k\rangle_2, \qquad j,k \in \{1, \dots, n\}.$$

We have

$$\langle \boldsymbol{e}_j, \boldsymbol{e}_k\rangle_2 = \boldsymbol{I}_n(j,k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k \end{cases}$$

and a direct calculation shows that

$$\langle \boldsymbol{U}\boldsymbol{e}_j, \boldsymbol{U}\boldsymbol{e}_k\rangle_2 = \sum_{i=1}^n \boldsymbol{U}(i,j)\bar{\boldsymbol{U}}(i,k) = (\boldsymbol{U}^T\bar{\boldsymbol{U}})(j,k).$$

Thus $\boldsymbol{U}^T\bar{\boldsymbol{U}} = \boldsymbol{I}_n$ which, upon conjugation, gives $\bar{\boldsymbol{U}}^T\boldsymbol{U} = \boldsymbol{I}_n$. From Theorem **??** this means that $\boldsymbol{U}$ is invertible with inverse $\bar{\boldsymbol{U}}^T$. This means that we also have $\boldsymbol{U}\bar{\boldsymbol{U}}^T = \boldsymbol{I}_n$.

(iv) $\implies$ (v) This was proved in the preceding part of the proof.

(v) $\implies$ (i) By hypothesis we have

$$\bar{\boldsymbol{U}}^T\boldsymbol{U} = \boldsymbol{I}_n \quad\implies\quad \boldsymbol{U}^{-1}\bar{\boldsymbol{U}}^{-T} = \boldsymbol{I}_n.$$

By Theorem **??** this implies that $\boldsymbol{U}^{-1}$ is invertible with inverse $\bar{\boldsymbol{U}}^{-T}$. Thus

$$\bar{\boldsymbol{U}}^{-T}\boldsymbol{U}^{-1} = \boldsymbol{I}_n \quad\implies\quad \boldsymbol{U}^{-T}\bar{\boldsymbol{U}}^{-1} = \boldsymbol{I}_n.$$

Let us define a basis $\{e'_1, \dots, e'_n\}$ for V by asking that

$$e'_{j_0} = \sum_{j=1}^n \boldsymbol{U}^{-1}(j, j_0)e_j. \tag{7.11}$$

The computation (7.9), but using $\boldsymbol{U}^{-1}$ in place of $\boldsymbol{U}$, gives

$$\langle e'_{j_1}, e'_{j_2}\rangle = \sum_{k=1}^n \boldsymbol{U}^{-1}(k, j_1)\bar{\boldsymbol{U}}^{-1}(k, j_2) = (\boldsymbol{U}^{-T}\bar{\boldsymbol{U}}^{-1})(j_1, j_2) = \boldsymbol{I}_n(j_1, j_2).$$

Thus

$$\langle e'_{j_1}, e'_{j_2}\rangle = \begin{cases} 1, & j_1 = j_2, \\ 0, & j_1 \neq j_2, \end{cases}$$

showing that $\{e'_1, \dots, e'_n\}$ is a Hilbert basis. Since (7.11) implies that

$$e_{j_0} = \sum_{j=1}^{n} \boldsymbol{U}(j, j_0) e'_j,$$

this part of the result follows. ■

In the case where $\mathbb{F} = \mathbb{R}$ the previous result, along with Theorem **??**, shows that the change of basis matrices between Hilbert bases are precisely the orthogonal matrices. The set of $n \times n$ orthogonal matrices were denoted by $\mathsf{O}(n)$. In the case where $\mathbb{F} = \mathbb{C}$ the matrices of the preceding result are called ***unitary*** matrices and the set of $n \times n$ unitary matrices are denoted by $\mathsf{U}(n)$.

One of the interesting features of Hilbert bases is that it is easy to determine the components of a vector relative to the basis. The following result records this.

7.3.17 Proposition (Components relative to a finite orthonormal set) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *let* $(\mathsf{V}, \langle\cdot,\cdot\rangle_2)$ *be a (not necessarily finite-dimensional)* $\mathbb{F}$*-inner product space, and let* $\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ *be a finite orthonormal set. If* $\mathrm{v} \in \mathrm{span}_{\mathbb{F}}(\mathrm{e}_1, \dots, \mathrm{e}_n)$ *then*

$$\mathrm{v} = \langle \mathrm{v}, \mathrm{e}_1\rangle \mathrm{e}_1 + \cdots + \langle \mathrm{v}, \mathrm{e}_n\rangle \mathrm{e}_n.$$

Proof This is Exercise 7.3.4. ■

The preceding result has the following obvious corollary.

7.3.18 Corollary (Components relative to a finite Hilbert basis) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *let* $(\mathsf{V}, \langle\cdot,\cdot\rangle_2)$ *be a finite-dimensional* $\mathbb{F}$*-inner product space, and let* $\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ *be a finite Hilbert basis for* V*. For* $\mathrm{v} \in \mathsf{V}$ *the components of* v *are* $\langle \mathrm{v}, \mathrm{e}_j\rangle$, $\mathrm{j} \in \{1, \dots, \mathrm{n}\}$.

We shall now give some properties of Hilbert bases for finite-dimensional inner product spaces that may, at first glance, seem obvious and/or silly. However, they arise in the infinite-dimensional setting in a rather less obvious and hopefully less silly way. Therefore, it is worth recording them in the present setup.

The first result is the finite-dimensional version of Bessel's inequality.

7.3.19 Proposition (Bessel's inequality for finite orthonormal sets) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is a (not necessarily finite-dimensional)* $\mathbb{F}$*-inner product space, and if* $\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ *is a finite orthonormal set, then, for any* $\mathrm{v} \in \mathsf{V}$,

$$\sum_{\mathrm{j}=1}^{\mathrm{n}} |\langle \mathrm{v}, \mathrm{e}_\mathrm{j}\rangle|^2 \leq \|\mathrm{v}\|^2.$$

Proof This is Exercise 7.3.5. ■

Our final result gives several conditions equivalent to that of being a Hilbert basis. These are more or less "obvious" in finite-dimensions, but are a little less so in infinite-dimensions.

7.3.20 Theorem (Characterisations of finite Hilbert bases) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a finite-dimensional $\mathbb{F}$-inner product space, and let $\mathscr{B} = \{\mathrm{e}_1, \dots, \mathrm{e}_\mathrm{n}\}$ be an orthonormal set. The following statements are equivalent:*

(i) *$\mathscr{B}$ is basic;*

(ii) *$\mathscr{B}$ is total;*

(iii) *for all $\mathrm{v} \in \mathsf{V}$ we have*

$$\|\mathrm{v}\|^2 = \sum_{\mathrm{j}=1}^{\mathrm{n}} |\langle \mathrm{v}, \mathrm{e}_\mathrm{j}\rangle|^2$$

*(**Parseval's equality**);*

(iv) *for all $\mathrm{u}, \mathrm{v} \in \mathsf{V}$ we have*

$$\langle \mathrm{u}, \mathrm{v}\rangle = \sum_{\mathrm{j}=1}^{\mathrm{n}} \langle \mathrm{u}, \mathrm{e}_\mathrm{j}\rangle\overline{\langle \mathrm{v}, \mathrm{e}_\mathrm{j}\rangle};$$

(v) *$\mathscr{B}^\perp = \{0_\mathsf{V}\}$;*

(vi) *$\mathscr{B}$ is a maximal.*

Proof We leave this to the reader as Exercise 7.3.6. ■

7.3.3 Countable orthonormal sets and countable Hilbert bases

In the finite-dimensional case we see that Hilbert bases are always bases in the usual sense. Thus a Hilbert basis for a finite-dimensional inner product space is simply an instance of something we are already familiar with. This is no longer true in infinite-dimensions. Complications can arise in multiple ways. From Theorem 7.3.12 we know that every inner product space possesses a maximal orthonormal subset. For Hilbert spaces, these maximal orthonormal sets are necessarily Hilbert bases by Corollary 7.3.13. However, in infinite-dimensions it is not necessarily the case that a Hilbert basis is a basis. It *can* be the case that a Hilbert basis is a basis (see Example 7.3.8–2), but it is also true that countable Hilbert bases for Hilbert spaces are *never* bases. Also, for non-Hilbert spaces it can happen that they do not possess a Hilbert basis (see Example 7.3.8–3).

What we do in this section is consider the special case of inner product spaces that admit a countable Hilbert basis. Thus we consider the case where we have a countable orthonormal set $(e_j)_{j\in\mathbb{Z}_{>0}}$ for an inner product space and we assume that for any $v \in \mathsf{V}$ we can write

$$v = \sum_{j=1}^{\infty} c_j e_j. \tag{7.12}$$

Note that this sum is infinite, not finite as for a Hamel basis. The definition of convergence we use for this sum is made exactly as with the discussion of series in Banach spaces in Definition 6.4.1. That is to say, the existence of the infinite sum in (7.12) means that, for every $\epsilon \in \mathbb{R}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that

$$\left\|v - \sum_{j=1}^{k} c_j e_j\right\| < \epsilon$$

for every $k \geq N$. Note that it is not obvious that this coincides with the notion of convergence used in our general discussion in Section 7.3.1. Indeed, convergence for series using general index sets as used in Section 7.3.1 is equivalent to unconditional convergence for series using the index set $\mathbb{Z}_{>0}$. This sort of convergence *implies* convergence in the usual sense, but is not equivalent to it. This notwithstanding, we shall see that the usual definition of convergence for series is the appropriate one to use in the setting of countable Hilbert bases.

First we establish the appropriate condition under which an inner product space admits a countable Hilbert basis. In Theorem 7.3.14 we saw that the appropriate condition for the existence of a finite Hilbert basis was that the inner product space be, not surprisingly, finite-dimensional. For countable Hilbert bases, the condition turns out to be that the inner product space be separable (see Definition 6.6.12).

7.3.21 Theorem (Characterisation of existence of countable Hilbert bases) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is a separable, infinite-dimensional $\mathbb{F}$-inner product space, then the following statements hold:*

(i) if V is a Hilbert space then it possesses a countably infinite Hilbert basis;

(ii) every Hilbert basis for V is countably infinite.

Proof By Corollary 7.3.13 we know that if V is a Hilbert space then it possesses a Hilbert basis and by Proposition 7.3.15 we know that every Hilbert basis is infinite. It remains to show that every Hilbert basis is countable. Suppose otherwise and so there exists an uncountable Hilbert basis $\mathscr{B} = \{e_i\}_{i\in I}$. If $i_1, i_2 \in I$ then

$$\|e_{i_1} - e_{i_2}\| = (\langle e_{i_1} - e_{i_2}, e_{i_1} - e_{i_2}\rangle)^{1/2} = (\|e_{i_1}\|^2 + \|e_{i_2}\|^2)^{1/2} = \sqrt{2}. \tag{7.13}$$

since e_{i_1} and e_{i_2} are orthogonal. For each $i \in I$ define $U_i = \mathsf{B}(\frac{1}{4}, e_i)$ and note that $U_{i_1} \cap U_{i_2} = \emptyset$ by (7.13). Now let $S \subseteq \mathsf{V}$ be countable. Then there exists an uncountable set $J \subseteq I$ such that $S \cap (\cup_{j\in J} U_j) = \emptyset$. Thus $S \subseteq \mathsf{V} \setminus (\cup_{j\in J} U_j)$ and so $\mathrm{cl}(S) \subseteq \mathsf{V} \setminus (\cup_{j\in J} U_j)$. Thus $\mathrm{cl}(S) \neq \mathsf{V}$ and so V is not separable. ■

The companion result to this is that countable Hilbert bases exist *only* for separable inner product spaces.

7.3.22 Theorem (Necessary conditions for a countable Hilbert basis) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is an $\mathbb{F}$-inner product space having a countably infinite Hilbert basis, then V is separable and infinite-dimensional.*

Proof That V is infinite-dimensional follows from Proposition 7.3.15. To show that V is separable we let $\mathscr{B} = \{e_j\}_{j\in\mathbb{Z}_{>0}}$ be a countably infinite Hilbert basis and let $\mathsf{V}_0 = \mathrm{span}_{\mathbb{F}}(\mathscr{B})$. By Theorem 7.3.9 we know that $\mathscr{B}$ is total and so $\mathrm{cl}(\mathsf{V}_0) = \mathsf{V}$. Now define

$$\mathbb{F}_{\mathbb{Q}} = \begin{cases} \mathbb{Q}, & \mathbb{F} = \mathbb{R}, \\ q_r + \mathrm{i}q_i, & \mathbb{F} = \mathbb{C} \end{cases}$$

and consider the set

$$S_{\mathscr{B}} = \{q_1 e_{j_1} + \cdots + q_k e_{j_k} \mid k \in \mathbb{Z}_{>0},\ q_1, \dots, q_k \in \mathbb{F}_{\mathbb{Q}}\}$$

of finite linear combinations of elements from $\mathscr{B}$ with coefficients in $\mathbb{F}_{\mathbb{Q}}$. Using Proposition **??** we may conclude that $S_{\mathscr{B}}$ is countable. We claim that $S_{\mathscr{B}}$ is dense in V. From Exercise 6.6.2 it suffices to show that $S_{\mathscr{B}}$ is dense in V_0. Let $v \in \mathsf{V}_0$ so that we may write

$$v = c_1 e_{j_1} + \cdots + c_k e_{j_k}$$

for some $j_1, \dots, j_k \in \mathbb{Z}_{>0}$ and $c_1, \dots, c_k \in \mathbb{F}$. Let $\epsilon \in \mathbb{R}_{>0}$ and choose $q_a \in \mathbb{F}_{\mathbb{Q}}$ such that $|c_a - q_a| < \frac{\epsilon}{k}$, $a \in \{1, \dots, k\}$. Then

$$\|v - q_1 e_{j_1} - \cdots - q_k e_{j_k}\| \le |c_1 - q_1|\|e_{j_1}\| + \cdots + |c_k - q_k|\|e_{j_k}\| < \epsilon$$

by the triangle inequality. Thus $v \in \operatorname{cl}(S_{\mathscr{B}})$ by Proposition 6.6.8. We have thus shown that the countable set $S_{\mathscr{B}}$ is dense in V, as desired. ■

First let us determine the form of the coefficients in the summation (7.12) if it does indeed converge. The reader should compare this result to Proposition 7.3.17.

7.3.23 Proposition (Components relative to a countable orthonormal set) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*-inner product space, and let* $\{\mathrm{e}_j\}_{j \in \mathbb{Z}_{>0}}$ *be an orthonormal set in* V*. If the sum*

$$\sum_{j=1}^{\infty} c_j e_j$$

converges to $\mathrm{v} \in \mathsf{V}$*, then for each* $j \in \mathbb{Z}_{>0}$*,* $c_j = \langle \mathrm{v}, \mathrm{e}_j \rangle$*.*

Proof By Proposition 7.2.1 and Theorem 6.5.2 we have

$$\langle v, e_k \rangle = \Big\langle \sum_{j=1}^{\infty} c_j e_j, e_k \Big\rangle = \Big\langle \lim_{n\to\infty} \sum_{j=1}^{n} c_j e_j, e_k \Big\rangle = \lim_{n\to\infty} \sum_{j=1}^{n} c_j \langle e_j, e_k \rangle = c_k$$

for every $k \in \mathbb{Z}_{>0}$. ■

The reader should be sure to appreciate that, while the formula for the coefficients is exactly as given in the finite-dimensional case in Proposition 7.3.17, one must be a little more careful in arriving at this formula as there are issues with swapping limits with the inner product that must be accounted for.

The following result holds even for orthonormal sets that are not basic and should be compared to Proposition 7.3.19.

7.3.24 Theorem (Bessel's inequality for countable orthonormal sets) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be a* $\mathbb{F}$*-inner product space, and let* $\mathscr{B} = \{\mathrm{e}_j\}_{j \in \mathbb{Z}_{>0}}$ *be a countably infinite orthonormal set. Then, for any* $\mathrm{v} \in \mathsf{V}$*, the sum*

$$\sum_{j=1}^{\infty} |\langle \mathrm{v}, \mathrm{e}_j \rangle|^2 \tag{7.14}$$

converges and satisfies

$$\sum_{j=1}^{\infty} |\langle \mathrm{v}, \mathrm{e}_j \rangle|^2 \le \|\mathrm{v}\|^2.$$

Proof Let v_k denote the kth partial sum:

$$v_k = \sum_{j=1}^{k} \langle v, e_j \rangle e_j.$$

We claim that for $j \in \{1, \dots, k\}$, e_j is orthogonal to $v - v_k$. Indeed,

$$\langle v - v_k, e_j \rangle = \langle v, e_j \rangle - \langle v_k, e_j \rangle.$$

We also have, by a direct computation, $\langle v_k, e_j \rangle$ as the jth term in the sum, i.e., $\langle v_k, e_j \rangle = \langle v, e_j \rangle$. Thus $\langle v - v_k, e_j \rangle = 0$ as claimed. From this, since v_k is a linear combination of $\{e_1, \dots, e_k\}$, it follows that $v - v_k$ and v_k are orthogonal. By the Pythagorean identity (Exercise 7.1.12) we then have

$$\|v\|^2 = \|v - v_k + v_k\|^2 = \|v - v_k\|^2 + \|v_k\|^2,$$

giving

$$\|v_k\|^2 \le \|v\|^2. \tag{7.15}$$

Since the vectors $\{e_1, \dots, e_k\}$ are orthonormal we compute

$$\|v_k\|^2 = \Big\langle \sum_{j=1}^{k} \langle v, e_j \rangle e_j, \sum_{l=1}^{k} \langle v, e_l \rangle e_l \Big\rangle = \sum_{j=1}^{k} \sum_{l=1}^{k} \langle v, e_j \rangle \overline{\langle v, e_l \rangle} \langle e_j, e_l \rangle = \sum_{j=1}^{k} |\langle v, e_j \rangle|^2. \tag{7.16}$$

Thus, combining (7.15) and (7.16), we have shown that the inequality

$$\sum_{j=1}^{k} |\langle v, e_j \rangle|^2 \le \|v\|^2$$

holds for any $k \in \mathbb{Z}_{>0}$. Thus the sum (7.14) is a sum of positive terms with each partial sum being bounded above by $\|v\|^2$. It follows that the sequence of partial sums must converge to a number being at most $\|v\|^2$. ■

We also have the following result which should be compared to Theorem 7.3.20.

7.3.25 Theorem (Characterisations of countable Hilbert bases) *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle \cdot, \cdot \rangle)$ be a separable $\mathbb{F}$-inner product space, and let $\mathscr{B} = \{\mathrm{e}_j\}_{j \in \mathbb{Z}_{>0}}$ be an orthonormal set. The following four statements are equivalent:*

(i) $\mathscr{B}$ is basic;

(ii) $\mathscr{B}$ is total;

(iii) for every $\mathrm{v} \in \mathsf{V}$ the equality

$$\|\mathrm{v}\|^2 = \sum_{j=1}^{\infty} |\langle \mathrm{v}, \mathrm{e}_j \rangle|^2$$

*holds (**Parseval's equality**);*

(iv) *for all* $\mathrm{u}, \mathrm{v} \in \mathsf{V}$ *we have*

$$\langle \mathrm{u}, \mathrm{v}\rangle = \sum_{\mathrm{j}=1}^{\infty} \langle \mathrm{u}, \mathrm{e_j}\rangle\overline{\langle \mathrm{v}, \mathrm{e_j}\rangle};$$

Also, the following two statements are equivalent:

(v) $\mathscr{B}^{\perp} = \{0_{\mathsf{V}}\}$*;*

(vi) $\mathscr{B}$ *is maximal.*

Finally, if V *is a Hilbert space, the first four equivalent statements are equivalent to the last two equivalent statements.*

Proof (i) $\implies$ (ii) Let $\mathscr{B} = \{e_j\}_{j\in\mathbb{Z}_{>0}}$ be basic and let $v \in \mathsf{V}$. We can then write

$$v = \sum_{j\in\mathbb{Z}_{>0}} c_j e_j$$

for some coefficients $c_j \in \mathbb{F}$, $j \in \mathbb{Z}_{>0}$. If we define

$$v_k = \sum_{j=1}^{k} c_j e_j$$

then the sequence $(v_k)_{k\in\mathbb{Z}_{>0}}$ converges to v. Thus $v \in \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(\mathscr{B}))$ and so $\mathscr{B}$ is total.

(ii) $\implies$ (iii) Let $v \in \mathsf{V}$. Since $\mathscr{B}$ is total there exists a sequence $(v_k)_{k\in\mathbb{Z}_{>0}}$ in $\mathrm{span}_{\mathbb{F}}(\mathscr{B})$ such that $v = \lim_{k\to\infty} v_k$. For each $k \in \mathbb{Z}_{>0}$ write

$$v_k = c_{k1}e_{j_{k1}} + \cdots + c_{km_k}e_{j_{km_k}}$$

for $m_k \in \mathbb{Z}_{>0}$, coefficients $c_{k1}, \dots, c_{km_k} \in \mathbb{F}$, and distinct $j_{k1}, \dots, j_{km_k} \in I$. By Proposition 7.3.23 it follows that $c_{kl} = \langle v_k, e_{j_{kl}}\rangle$ for each $k \in \mathbb{Z}_{>0}$, $l \in \{1, \dots, m_k\}$. This means that we can write

$$v_k = \sum_{j=1}^{\infty} \langle v_k, e_j\rangle e_j$$

for each $k \in \mathbb{Z}_{>0}$, with the sum being finite.

We may also directly compute (cf. the proof of Theorem 7.3.24)

$$\|v_k\|^2 = \sum_{j=1}^{\infty} |\langle v_k, e_j\rangle|^2,$$

using the fact that the inner product commutes with finite sums. Now, using continuity of the norm and inner product, along with Theorem 6.5.2, gives

$$\|v\|^2 = \lim_{k\to\infty} \|v_k\|^2 = \lim_{k\to\infty} \sum_{j=1}^{\infty} |\langle v_k, e_j\rangle|^2 = \sum_{j=1}^{\infty} |\langle v, e_j\rangle|^2,$$

as desired.

(iii) $\implies$ (iv) For $u, v \in \mathsf{V}$ we have

$$\begin{aligned}
&\|u+v\|^2 = \sum_{j=1}^{\infty} |\langle u+v, e_j\rangle|^2 \\
\implies \quad &\|u\|^2 + \|v\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} \\
&= \sum_{j=1}^{\infty} |\langle u, e_j\rangle|^2 + \sum_{j=1}^{\infty} |\langle v, e_j\rangle|^2 + \sum_{j=1}^{\infty} (\langle u, e_j\rangle\overline{\langle v, e_j\rangle} + \overline{\langle u, e_j\rangle\overline{\langle v, e_j\rangle}}) \\
\implies \quad &\operatorname{Re}(\langle u, v\rangle) = \sum_{j=1}^{\infty} \operatorname{Re}(\langle u, e_j\rangle\overline{\langle v, e_j\rangle}).
\end{aligned}$$

If $\mathbb{F} = \mathbb{R}$ this establishes the result. If $\mathbb{F} = \mathbb{C}$, a similar computation using the equality

$$\|u + \mathrm{i}v\|^2 = \sum_{j=1}^{\infty} |\langle u + \mathrm{i}v, e_j\rangle|^2$$

gives

$$\operatorname{Im}(\langle u, v\rangle) = \sum_{j=1}^{\infty} \operatorname{Im}(\langle u, e_j\rangle\overline{\langle v, e_j\rangle}).$$

(iv) $\implies$ (i) Since part (iv) obviously implies part (iii), we shall prove that (iii) implies (i). Thus we have

$$\|v\|^2 = \sum_{j=1}^{\infty} |\langle v, e_j\rangle|^2$$

for every $v \in \mathsf{V}$. For $k \in \mathbb{Z}_{>0}$ let us define

$$v_k = \sum_{j=1}^{k} \langle v, e_j\rangle e_j.$$

Note that

$$\begin{aligned}
\langle v - v_k, v_k\rangle &= \Big\langle v - \sum_{j=1}^{k} \langle v, e_j\rangle e_j, \sum_{l=1}^{k} \langle v, e_l\rangle e_l \Big\rangle \\
&= \Big\langle v, \sum_{l=1}^{k} \langle v, e_l\rangle e_l \Big\rangle - \Big\langle \sum_{j=1}^{k} \langle v, e_j\rangle e_j, \sum_{l=1}^{k} \langle v, e_l\rangle e_l \Big\rangle \\
&= \sum_{l=1}^{k} |\langle v, e_l\rangle|^2 - \sum_{j=1}^{k} |\langle v, e_j\rangle|^2 = 0
\end{aligned}$$

for every $k \in \mathbb{Z}_{>0}$. By the Pythagorean equality,

$$\|v\|^2 = \|v - v_k + v_k\|^2 = \|v - v_k\|^2 + \|v_k\|^2 \quad \implies \quad \|v - v_k\|^2 = \|v\|^2 - \|v_k\|^2.$$

By assumption,

$$\lim_{k\to\infty} \|v_k\|^2 = \|v\|^2$$

and so

$$\lim_{k\to\infty}\|v - v_k\| = 0,$$

implying that

$$v = \sum_{j=1}^{\infty}\langle v, e_j\rangle e_j,$$

and so in particular implying that $\mathscr{B}$ is basic.

(v) $\implies$ (vi) Suppose that $\mathscr{B}$ is not maximal. Then there exists an orthonormal set $\mathscr{B}'$ such that $\mathscr{B} \subset \mathscr{B}'$. Let $v \in \mathscr{B}' \setminus \mathscr{B}$. Then, clearly, $v \in \mathscr{B}^{\perp}$ and $v \neq 0_{\mathsf{V}}$. Thus $\mathscr{B}^{\perp} \neq \{0_{\mathsf{V}}\}$.

(vi) $\implies$ (v) Suppose that $\mathscr{B}^{\perp} \neq \{0_{\mathsf{V}}\}$ and let $v \in \mathscr{B}^{\perp}$ have unit length. Then the set $\mathscr{B} \cup \{v\}$ is an orthonormal set that strictly contains $\mathscr{B}$. Thus $\mathscr{B}$ is not maximal.

(ii) $\implies$ (v) By Proposition 7.1.13(iv) we have $\mathscr{B}^{\perp} = \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(\mathscr{B}))^{\perp}$. From this fact, if $\mathscr{B}$ is total it immediately follows that $\mathscr{B}^{\perp} = \{0_{\mathsf{V}}\}$.

(vi) $\implies$ (i) (assuming V is a Hilbert space) Let $v \in \mathsf{V}$. Bessel's inequality gives

$$\sum_{j=1}^{\infty}|\langle v, e_j\rangle|^2 \leq \|v\|^2,$$

and this implies that the series on the right converges and so is Cauchy. Let $\epsilon \in \mathbb{R}_{>0}$ and let $N \in \mathbb{Z}_{>0}$ be such that

$$\sum_{j=k+1}^{l}|\langle v, e_j\rangle|^2 < \epsilon$$

for every $k, l \geq N$ with $l > k$. A direct computation using properties of inner products then gives

$$\Big\|\sum_{j=k+1}^{l}\langle v, e_j\rangle e_j\Big\|^2 = \sum_{j=k+1}^{l}|\langle v, e_j\rangle|^2 < \epsilon,$$

which shows that the series

$$\sum_{j=1}^{\infty}\langle v, e_j\rangle e_j$$

is Cauchy. By Theorem 6.4.17 this series converges, implying that $\mathscr{B}$ is basic. ■

7.3.4 Generalised Fourier series

In our general framework, the notion of a Fourier series is easily discussed. We shall discuss Fourier series (although we will think of this as being a means of getting at the inverse of the so-called CDFT) in Chapter 12. In this case, as we shall see, other issues not present in our general inner product space constructions, become relevant. Thus we focus our discussion in this section on the generalities. This will allow us to separate out these general considerations from the more specific ones in Chapter 12.

We begin with a definition that at this point is simply the giving of a name to something we already have been talking about.

7.3.26 Definition (Generalised Fourier series) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a $\mathbb{F}$-inner product space. If $\{e_i\}_{i\in I}$ is a Hilbert basis for V and if $v \in \mathsf{V}$, the ***generalised Fourier series*** for v is the series

$$v = \sum_{i\in I} \langle v, e_i\rangle e_i,$$

which converges to v. •

Let us consider some general and, therefore, more or less elementary examples.

7.3.27 Examples (Generalised Fourier series)

1. Let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a finite-dimensional inner product space with Hilbert basis $\{e_1, \dots, e_n\}$. The generalised Fourier series for $v \in \mathsf{V}$ is then simply the representation of v in the (Hamel) basis $\{e_1, \dots, e_n\}$, just as prescribed by Corollary 7.3.18:

$$v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n.$$

2. Next consider the inner product space $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_1)$ with its standard basis $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$; this is a Hilbert basis as we saw in Example 7.3.8–2. In this case the Hilbert basis is also a basis in the usual sense. Thus the generalised Fourier series for $\boldsymbol{v} \in \mathbb{F}_0^\infty$,

$$\boldsymbol{v} = \sum_{j=1}^\infty \boldsymbol{v}(j)\boldsymbol{e}_j,$$

is simply the representation of $\boldsymbol{v}$ with respect to a basis in the usual sense.

3. Finally, let us consider the completion $(\ell^2(\mathbb{F}), \langle\cdot,\cdot\rangle_2)$ of $(\mathbb{F}_0^\infty, \langle\cdot,\cdot\rangle_2)$. In this case the generalised Fourier series for $\boldsymbol{v} \in \ell^2(\mathbb{F})$ has the form

$$\boldsymbol{v} = \sum_{j=1}^\infty \boldsymbol{v}(j)\boldsymbol{e}_j.$$

Note that this is *not* the representation of $\boldsymbol{v}$ in a basis in the usual sense because the sum is possibly finite. Indeed, it is quite clear that $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ is not a (Hamel;) basis. Moreover, we shall see in Theorem 7.3.36 that any (Hamel) basis for $\ell^2(\mathbb{F})$ has cardinality strictly greater than that of $\mathbb{Z}_{>0}$. •

The preceding two examples illustrate the difference between the purely algebraic notion of a Hamel basis and the analytical notion of a Hilbert basis. It is probably worth understanding the message these examples are trying to pass on.

Let us now give a useful geometric interpretation of the generalised Fourier series. We recall from Section 7.1.5 the notation $\operatorname{dist}(v, S)$ for the distance from $v \in \mathsf{V}$ to a subset $S \subseteq \mathsf{V}$.

7.3.28 Theorem (The best approximation property of generalised Fourier series) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*-inner product space, and let* $\{e_i\}_{i\in I}$ *be a Hilbert basis for* V*. For* $J \subseteq I$ *let us abbreviate*

$$\mathsf{V}_J = \mathrm{cl}(\mathrm{span}_{\mathbb{F}}(e_j|\ j \in J)),$$

and assume that V_J *is complete. If* $v \in \mathsf{V}$ *and if* $J \subseteq I$*, then*

$$v_J \triangleq \sum_{j\in J} \langle v, e_j\rangle e_j$$

is the unique vector in V_J *for which* $\mathrm{dist}(v, \mathsf{V}_J) = \|v - v_J\|$.

Proof We first claim that the series v_J converges. Since we are assuming that V_J is complete, it suffices by Theorem 6.4.17 to show that the series v_J is Cauchy. Let $\epsilon \in \mathbb{R}_{>0}$. Since the series

$$\sum_{j\in J} |\langle v, c_j\rangle|^2$$

is convergent by Theorem 7.3.6 it is also Cauchy. Thus there exists a finite subset $J \subseteq I$ such that

$$\sum_{j\in J'} |\langle v, e_j\rangle|^2 < \epsilon$$

for every finite subset $J' \subseteq I$ for which $J' \cap J = \emptyset$. Then, by Theorem 7.3.20,

$$\Big\|\sum_{j\in J'} \langle v, e_j\rangle e_j\Big\|^2 = \sum_{j\in J'} |c_j|^2 < \epsilon$$

for every finite subset $J' \subseteq I$ for which $J' \cap J = \emptyset$. This gives convergence of the series for v_J, as desired.

Now, by Theorem 7.1.26, it suffices to show that $v - v_J \in \mathsf{V}_J^{\perp}$. By Proposition 7.1.13(iv) it suffices to show that $\langle v - v_J, e_j\rangle = 0$ for every $j \in J$. But this holds since

$$\langle v - v_J, e_j\rangle = \Big\langle v - \sum_{j'\in J} \langle v, e_{j'}\rangle e_{j'}, e_j\Big\rangle = \langle v, e_{j'}\rangle - \langle v, e_{j'}\rangle = 0,$$

where we swap the sum and inner product by Proposition 7.2.1 and Theorem 6.5.2. ■

The preceding discussion has to do with representing a vector in an inner product space by a generalised Fourier series. The next result tells us that any “reasonable” collection of coefficients are those of a generalised Fourier series.

7.3.29 Theorem (Riesz–Fischer[5] Theorem) *Let* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *let* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *be an* $\mathbb{F}$*-Hilbert space, and let* $(e_i)_{i\in I}$ *be an orthonormal family. If* $(c_i)_{i\in I}$ *is a family of numbers such that the series*

$$\sum_{i\in I} |c_i|^2 \tag{7.17}$$

[5] Frigyes Riesz (1880–1956) was born in what is now Hungary and was one of the founders of functional analysis. His younger brother Marcel was also a mathematician of some note. Ernst Sigismund Fischer (1875–1954) was an Austrian mathematician whose contributions to mathematics were in the areas of algebra and analysis.

converges in the sense of Definition 2.4.31, then the series

$$\sum_{\mathrm{i}\in\mathrm{I}} \mathrm{c_i e_i} \tag{7.18}$$

converges in the sense of Definition 6.4.16. Moreover, if the series converges to $\mathrm{v} \in \mathsf{V}$ *then* $\mathrm{c_i} = \langle \mathrm{v}, \mathrm{e_i}\rangle$, $\mathrm{i} \in \mathrm{I}$.

Proof We claim that the sum (7.18) is Cauchy. Let $\epsilon \in \mathbb{R}_{>0}$. Since the series (7.17) is convergent and so Cauchy, there exists a finite set $J \subseteq I$ such that

$$\sum_{j\in J'} |c_j|^2 < \epsilon$$

for every finite subset $J' \subseteq I$ for which $J \cap J' = \emptyset$. By Theorem 7.3.20 we then have

$$\Big\|\sum_{j\in J'} c_j e_j\Big\|^2 = \sum_{j\in J'} |c_j|^2 < \epsilon$$

for every finite subset $J' \subseteq I$ for which $J \cap J' = \emptyset$. Thus the series (7.18) is Cauchy, and so convergent by Theorem 6.4.17. The last assertion is simply Proposition 7.3.5. ■

7.3.5 Classification of Hilbert spaces

In this section we use the idea of Hilbert bases to characterise all Hilbert spaces. As we shall see, the classification is actually quite simple, just as with the classification of all vector spaces induced by the size of their bases.

First let us assert that the dimension of an inner product space, when it exists, is well defined.

7.3.30 Theorem (Invariance of cardinality of maximal orthonormal sets) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *if* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is an* $\mathbb{F}$*-inner product space, and if* $\{\mathrm{e_i}\}_{\mathrm{i}\in\mathrm{I}}$ *and* $\{\mathrm{f_j}\}_{\mathrm{j}\in\mathrm{J}}$ *are Hilbert bases for* V*, then* $\mathrm{card}(\mathrm{I}) = \mathrm{card}(\mathrm{J})$.

Proof If V possesses a finite maximal orthonormal set, then this set is a Hilbert basis and so also a Hamel basis by Theorem 7.3.14. Moreover, from the same result, every Hilbert basis for V is a Hamel basis. By Theorem 4.3.25 every Hamel basis for V has the same cardinality, and so the result follows when V has a finite maximal orthonormal set.

Next suppose that V has two infinite maximal orthonormal sets $\{e_i\}_{i\in I}$ and $\{f_j\}_{j\in J}$. For $j \in J$ denote

$$I_j = \{i \in I \mid \langle f_j, e_i\rangle \neq 0\}.$$

Since, by Theorem 7.3.6, we have

$$\sum_{i\in I} \langle f_j, e_i\rangle \leq \|f_j\|^2 = 1,$$

it follows from Proposition 2.4.33 that I_j is countable for each $j \in J$. We claim that $I = \cup_{j\in J} I_j$. It is clear that $\cup_{j\in J} I_j \subseteq I$. Suppose that the converse inclusion does not hold and let $i \in I \setminus (\cup_{j\in J} I_j)$. This means, by definition of the sets I_j, $j \in J$, that $\langle f_j, e_i\rangle = 0$

for every $j \in J$. By Theorem 7.3.9 this means that $f_j = 0_{\mathsf{V}}$; from this we conclude that $I \subseteq \cup_{j \in J} I_j$. Now we have

$$\operatorname{card}(I) = \operatorname{card}(\cup_{j \in J} I_j) \leq \operatorname{card}(\mathbb{Z}_{>0}) \operatorname{card}(J) \leq \operatorname{card}(J) \operatorname{card}(J) = \operatorname{card}(J),$$

using Theorem **??** and its Corollary **??**. By swapping the rôles of I and J we similarly prove that $\operatorname{card}(J) \leq \operatorname{card}(I)$, and so the theorem follows from Theorem **??**. ■

The result has the following obvious (by Theorem 7.3.9) corollary.

7.3.31 Corollary (Invariance of cardinality of Hilbert bases) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, if $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ is an $\mathbb{F}$-inner product space, and if $\{\mathrm{e_i}\}_{\mathrm{i} \in \mathrm{I}}$ and $\{\mathrm{f_j}\}_{\mathrm{j} \in \mathrm{J}}$ are Hilbert bases for V, then* $\operatorname{card}(\mathrm{I}) = \operatorname{card}(\mathrm{J})$.

The preceding theorem and corollary make sense of the following definition.

7.3.32 Definition (Hilbert dimension) Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be a $\mathbb{F}$-inner product space. The ***Hilbert dimension*** of V is the cardinality of any maximal orthonormal set in V. We denote by $\operatorname{hdim}_{\mathbb{F}}(\mathsf{V})$ the Hilbert dimension of V. •

For vector spaces we saw in Proposition 4.3.30 that the dimension was an isomorphism invariant, indeed the only isomorphism. That is to say, two vector spaces are isomorphic if and only if they have the same dimension. We would like to establish a similar assertion for inner product spaces, but replacing "dimension" with "Hilbert dimension" and replacing "isomorphism" with "isomorphism of inner product spaces." But such a result is not actually true, as we shall see. The desired result *is* true, however, if we restrict ourselves to the most interesting case of Hilbert spaces.

7.3.33 Theorem (Hilbert dimension characterises Hilbert spaces) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and if $(\mathsf{V}_1, \langle\cdot,\cdot\rangle_1)$ and $(\mathsf{V}_2, \langle\cdot,\cdot\rangle_2)$ are $\mathbb{F}$-Hilbert spaces, then the following statements are equivalent:*

(i) V_1 *and* V_2 *are isomorphic as inner product spaces;*

(ii) $\operatorname{hdim}_{\mathbb{F}}(\mathsf{V}_1) = \operatorname{hdim}_{\mathbb{F}}(\mathsf{V}_2)$.

Proof (i) $\Longrightarrow$ (ii) Let $\mathsf{L} \colon \mathsf{V}_1 \to \mathsf{V}_2$ be an inner product space isomorphism and let $\mathscr{B}_1$ be a Hilbert basis for V_1. Define

$$\mathscr{B}_2 = \{\mathsf{L}(u) \mid u \in \mathscr{B}_1\};$$

we claim that $\mathscr{B}_2$ is a Hilbert basis for V_2. First let us prove that $\mathscr{B}_2$ is orthonormal. If $\mathsf{L}(u_1), \mathsf{L}(u_2) \in \mathscr{B}_2$ we have

$$\langle \mathsf{L}(u_1), \mathsf{L}(u_2) \rangle_2 = \langle u_1, u_2 \rangle_2,$$

using the fact that L is an isomorphism of inner product spaces. Thus $\mathsf{L}(u_1)$ and $\mathsf{L}(u_2)$ are orthogonal if and only if they are distinct. Similarly one computes $\|\mathsf{L}(u)\| = 1$ for $u \in \mathscr{B}_1$. Thus $\mathscr{B}_2$ is indeed orthonormal. Now suppose that $v_0 \in \mathscr{B}_2^{\perp}$ and let $u_0 = \mathsf{L}^{-1}(v)$. Then, for every $u \in \mathscr{B}_1$,

$$\langle v_0, \mathsf{L}(u) \rangle_2 = \langle u_0, u \rangle_1 = 0,$$

implying that $u_0 = 0_{\mathsf{V}_1}$ by Theorem 7.3.9, since L is an isomorphism of inner product spaces, and since $\mathscr{B}_1$ is maximal. We conclude that $\mathscr{B}_2$ is maximal and so a Hilbert basis By Theorem 7.3.9.

(ii) $\Longrightarrow$ (i) Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be Hilbert bases for V_1 and V_2, respectively. By assumption there exists a bijection $\phi\colon \mathscr{B}_1 \to \mathscr{B}_2$. Note that by Theorem 7.3.9 every vector in V_1 can be written as

$$\sum_{u\in\mathscr{B}_1} c_u u$$

for coefficients $c_u \in \mathbb{F}$, $u \in \mathscr{B}_1$, such that

$$\sum_{u\in\mathscr{B}_1} |c_u|^2 < \infty.$$

Using this fact, let us define $\mathsf{L}\colon \mathsf{V}_1 \to \mathsf{V}_2$ by

$$\mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} c_u u\Big) = \sum_{u\in\mathscr{B}_1} c_u \phi(u).$$

We must show that L is well-defined and is an isomorphism of inner product spaces. To show that L is well-defined, we must show that it defines an element of V_2. This, however, follows from the Riesz-Fischer Theorem. Linearity of L follows from the fact that $\mathsf{L}(u) = \phi(u)$ for every $u \in \mathscr{B}_1$ (why?) and from the calculations

$$\mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} (a_u u + b_u u)\Big) = \sum_{u\in\mathscr{B}} a_u\phi(u) + \sum_{u\in\mathscr{B}} b_u\phi(u) = \mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} a_u u\Big) + \mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} b_u u\Big),$$

for $a_u, b_u \in \mathbb{F}$, $u \in \mathscr{B}_1$, and

$$\mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} \alpha(c_u u)\Big) = \alpha \sum_{u\in\mathscr{B}_1} c_u\phi(u) = \alpha\mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} c_u u\Big),$$

for $\alpha \in \mathbb{F}$ and $c_u \in \mathbb{F}$, $u \in \mathscr{B}_1$. (Of course, in the above computations we require that $\sum_{u\in\mathscr{B}_1} |a_u|^2$, $\sum_{u\in\mathscr{B}_1} |b_u|^2$, and $\sum_{u\in\mathscr{B}_1} |c_u|^2$ be finite.) The swapping of sums with addition and multiplication is justified by Proposition 7.2.1 and Theorem 6.5.2. Finally, we must show that L preserves the inner product. Using Theorem 7.3.9 we compute

$$\begin{aligned}\Big\langle \mathsf{L}\Big(\sum_{u\in\mathscr{B}_1} a_u u\Big), \mathsf{L}\Big(\sum_{u'\in\mathscr{B}_1} b_{u'} u'\Big)\Big\rangle_2 &= \Big\langle \sum_{u\in\mathscr{B}_1} a_u\phi(u), \sum_{u'\in\mathscr{B}_1} b_{u'}\phi(u')\Big\rangle_2 \\ &= \sum_{u\in\mathscr{B}_1} a_u\overline{b_u} \\ &= \Big\langle \sum_{u\in\mathscr{B}_1} a_u u, \sum_{u'\in\mathscr{B}_1} b_{u'} u'\Big\rangle_1,\end{aligned}$$

as desired. ■

Now that we have decided that the Hilbert dimension of a Hilbert space is its only property invariant under isomorphism of inner product spaces, let us provide for the set of Hilbert spaces with a prescribed Hilbert dimension a simple representative. It is perhaps useful to remind ourselves how this is done for vector spaces. If V is a vector space over a field F with dimension $\mathrm{card}(I)$, then we showed in Theorem 4.3.46 that V is isomorphic to the direct sum $\bigoplus_{i\in I}\mathsf{F}$. Thus the direct sum $\bigoplus_{i\in I}\mathsf{F}$ serves as a simple representative of *all* vector spaces with dimension equal to V. The situation is rather similar for Hilbert spaces.

The following theorem describes the simple representative we are after.

7.3.34 Theorem (A "canonical" Hilbert space of a prescribed Hilbert dimension) *For* $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$ *and for a set* I, *define*

$$\ell^2(\mathrm{I};\mathbb{F})=\Big\{\phi\colon\mathrm{I}\to\mathbb{F}\;\Big|\;\sum_{\mathrm{i}\in\mathrm{I}}|\phi(\mathrm{i})|^2<\infty\Big\}$$

and define an inner product on $\ell^2(\mathrm{I};\mathbb{F})$ *by*

$$\langle\phi,\psi\rangle_2=\sum_{\mathrm{i}\in\mathrm{I}}\phi(\mathrm{i})\overline{\psi(\mathrm{i})}.$$

Then $(\ell^2(\mathrm{I};\mathbb{F}),\langle\cdot,\cdot\rangle_2)$ *is a Hilbert space with Hilbert dimension* card(I).

Proof Note that $\ell^2(I;\mathbb{F})=\ell^2(\bigoplus_{i\in I}\mathbb{F})$ in the context of Definition 6.7.26. It then follows from Theorem 6.7.27 that $\ell^2(I;\mathbb{F})$ is a Banach space with respect to the norm $\|\cdot\|_2$ defined by

$$\|\phi\|_2=\sum_{i\in I}|\phi(i)|^2.$$

In order to show that it is a Hilbert space we should show that the norm is derived from the given inner product $\langle\cdot,\cdot\rangle_2$. First of all, for $\phi,\psi\in\ell^2(I;\mathbb{F})$, by Proposition 2.4.33 there exists an injection $\kappa\colon\mathbb{Z}_{>0}\to I$ such that $\phi(i)=\psi(i)=0$ for $i\notin\mathrm{image}(\kappa)$ and such that

$$\sum_{i\in I}|\phi(i)|^2=\sum_{j=1}^{\infty}|\phi(\kappa(j))|^2,\quad\sum_{i\in I}|\psi(i)|^2=\sum_{j=1}^{\infty}|\psi(\kappa(j))|^2.$$

Then, for $n\in\mathbb{Z}_{>0}$,

$$\Big|\sum_{j=1}^{n}\phi(\kappa(j))\overline{\psi(\kappa(j))}\Big|\le\Big(\sum_{j=1}^{n}|\phi(\kappa(j))|^2\Big)^{1/2}\Big(\sum_{j=1}^{n}|\psi(\kappa(j))|^2\Big)^{1/2},$$

using the Cauchy–Bunyakovsky–Schwarz inequality. Letting $n\to\infty$ we get

$$\Big|\sum_{i\in I}\phi(i)\overline{\psi(i)}\Big|^2\le\|\phi\|_2\|\psi\|_2<\infty.$$

Thus the sum defining the inner product converges. Completing the proof is now a matter of verifying the inner product axioms for $\langle\cdot,\cdot\rangle_2$, justifying the swapping of infinite sums and inner products using Proposition 7.2.1 and Theorem 6.5.2. ■

From the preceding result and from Theorem 7.3.33 (and its proof) we deduce the following interesting conclusion.

7.3.35 Corollary (Characterisation of Hilbert spaces up to isomorphism of inner product spaces) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, *if* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is a Hilbert space, and if* $\{e_i\}_{i\in I}$ *is a Hilbert basis for* V, *then the map* $\mathsf{L}\colon \mathsf{V} \to \ell^2(I;\mathbb{F})$ *defined by*

$$\mathsf{L}\Big(\sum_{i\in I} c_i e_i\Big) = \sum_{i\in I} c_i \boldsymbol{e}_i$$

is an isomorphism of inner product spaces.

It is worth digesting and understanding clearly the difference between the preceding corollary and its counterpart Theorem 4.3.46 for vector spaces. For a given set I the "canonical" $\mathbb{F}$-vector space of Hamel dimension card(I) is $\mathbb{F}_0^I$ and the "canonical" $\mathbb{F}$-Hilbert space of Hilbert dimension card(I) is $\ell^2(I;\mathbb{F})$. Both are subspaces of $\mathbb{F}^I$ (see Notation 4.3.45 for this notation). Moreover, $\mathbb{F}_0^I$ is a subspace of $\ell^2(I;\mathbb{F})$, and is a strict subspace unless card(I) is finite. Indeed, $\mathbb{F}_0^I$ and $\ell^2(I;\mathbb{F})$ are rather different objects when card(I) is not finite. For example, to make sense of the vector space $\ell^2(I;\mathbb{F})$ requires some analysis that is not required to make sense of $\mathbb{F}_0^I$. Note, for example, that we have not defined $\ell^2(I;\mathsf{F})$ for a general field F as a general field does not possess the absolute value structure of $\mathbb{R}$ or $\mathbb{C}$ that is needed to make things go. Thus $\ell^2(I;\mathbb{F})$ is, in some sense, a "deeper" object than $\mathbb{F}_0^I$. However, there is a strong connection between $\mathbb{F}_0^I$ and $\ell^2(I;\mathbb{F})$ in that the latter is the completion of the former if one uses the inner product

$$\langle\phi,\psi\rangle_2 = \sum_{i\in I} \phi(i)\overline{\psi(i)} \quad \text{(sum finite)}$$

on $\mathbb{F}_0^I$.

The preceding discussion leads one to the following natural question: "What is the relationship between the Hamel dimension and the Hilbert dimension of an inner product space?" For inner product spaces the answer can be, "They are equal." For example, $\mathbb{F}_0^I$ with the inner product $\langle\cdot,\cdot\rangle_2$ defined above has the same Hilbert and Hamel dimension. The question is deeper for Hilbert spaces. Indeed, from Theorem 6.6.26 and Theorem 7.3.21 we have the following result.

7.3.36 Theorem (Dimension of separable Hilbert space) *If* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *and if* $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ *is a separable infinite-dimensional* $\mathbb{F}$*-Hilbert space, then* $\dim_{\mathbb{F}}(\mathsf{V}) = \mathrm{card}(\mathbb{R})$.

In particular, this shows that $\mathbb{F}_0^\infty$ and $\ell^2(\mathbb{F})$ have different Hamel dimension, and so are not isomorphic. The story for Hilbert spaces of general dimension is more complicated, and we refer to the notes in Section 7.3.6.

7.3.6 Notes

The Riesz–Fischer Theorem was published independently by **ESF:07** and **FR:07a**, **FR:07b**.

JWE/RAT:70 study the relationship between the Hamel and Schauder dimensions of a Banach space. Applying their result to Hilbert spaces, their conclusions are that there is a condition on the cardinal numbers that characterise those infinite-dimensional Hilbert spaces whose Hamel and Hilbert dimensions agree. They point out that $\aleph_0 = \operatorname{card}(\mathbb{Z}_{>0})$ does not satisfy this condition (and so separable Hilbert spaces necessarily have different Hamel and Hilbert dimension) while $\aleph_1 = \operatorname{card}(\mathbb{R})$ does satisfy this condition (and so the Hilbert space $\ell^2(\mathbb{R};\mathbb{F})$ has equal Hamel and Hilbert dimension). The proof of **JWE/RAT:70** assumes the so-called Generalised Continuum Hypothesis which asserts that $2^{\aleph_o} = \aleph_{o+1}$ for every ordinal o.[6]

Exercises

7.3.1 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space. Show that an orthogonal set is linearly independent.

7.3.2 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Show that the standard basis $\{\boldsymbol{e}_j\}_{j\in\mathbb{Z}_{>0}}$ for $\mathbb{F}_0^\infty$ is a maximal orthonormal family.

7.3.3 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $(\mathsf{V}, \langle\cdot,\cdot\rangle)$ be an $\mathbb{F}$-inner product space, and let J be either the set $\{1, \dots, n\}$ for some $n \in \mathbb{Z}_{>0}$ or the set $\mathbb{Z}_{>0}$. Show that if $(u_j)_{j\in J}$ is orthonormal, then applying the Gram–Schmidt orthonormalisation procedure to this family gives the same family back again.

7.3.4 Prove Proposition 7.3.17. Point out the parts of your argument that are not valid in the infinite-dimensional case.

7.3.5 Prove Proposition 7.3.19. Point out the parts of your argument that are not generally valid for countable orthonormal sets $(e_j)_{j\in\mathbb{Z}_{>0}}$.

7.3.6 Prove Theorem 7.3.20. Point out the parts of your argument that are not generally valid for countable orthonormal sets $(e_j)_{j\in\mathbb{Z}_{>0}}$.

In the following exercise you will see just how fine is the notion of a maximal orthonormal set. Taking away any vector, or attempting to add a vector, ruins the maximality.

7.3.7 Let $\mathscr{B} = \{e_j\}_{j\in\mathbb{Z}_{>0}}$ be a maximal orthonormal set in an inner product space $(\mathsf{V}, \langle\cdot,\cdot\rangle)$.

(a) Show that for any $k \in \mathbb{Z}_{>0}$ the set $\mathscr{B} \setminus \{e_k\}$ is not maximal.

(b) Show that there is no vector $e_0 \in \mathsf{V}$ with the property that $\{e_0\} \cup \mathscr{B}$ is a maximal orthonormal set.

7.3.8 Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and let $(\mathsf{V}, \langle\cdot,\cdot\rangle_2)$ be an $\mathbb{F}$-Hilbert space. Let $\{e_i\}_{i\in I}$ be a Hilbert basis. Show that if $\alpha \in \mathsf{V}^*$ then the vector $v_\alpha \in \mathsf{V}$ associated with α by Corollary 7.2.5 satisfies $\langle v_\alpha, e_i\rangle = \overline{\alpha(e_i)}$ for each $i \in I$.

[6] The cardinals $\aleph_o$, defined for ordinals o, are defined using transfinite recursion as follows. Take $\aleph_0$ to be the cardinality of $\mathbb{Z}_{>0}$. Assuming that $\aleph_o$ has been defined, one defines $\aleph_{o+1}$ to be the successor (see Definition **??**) of $\aleph_o$.