

The physical foundations of geometric mechanics

Andrew D. Lewis*

2015/11/06

Last updated: 2017/09/28

Abstract

The principles of geometric mechanics are extended to the physical elements of mechanics, including space and time, rigid bodies, constraints, forces, and dynamics. What is arrived at is a comprehensive and rigorous presentation of basic mechanics, starting with precise formulations of the physical axioms. A few components of the presentation are novel. One is a mathematical presentation of force and torque, providing certain well-known, but seldom clearly explicated, fundamental theorems about force and torque. The classical principles of Virtual Work and Lagrange–d’Alembert are also given clear mathematical statements in various guises and contexts. Another novel facet of the presentation is its derivation of the Euler–Lagrange equations. Standard derivations of the Euler–Lagrange equations from the equations of motion for Newtonian mechanics are typically done for interconnections of particles. Here this is carried out in a coordinate-free manner for rigid bodies, giving for the first time a direct geometric path from the Newton–Euler equations to the Euler–Lagrange equations in the rigid body setting.

Keywords. Rigid body dynamics, force and torque, Newton–Euler equations, Euler–Lagrange equations, nonholonomic constraints

AMS Subject Classifications (2020). 70E55, 70G45, 70G65, 70G75, 70H03

Contents

1. Introduction	3
1.1 Features of the presentation.	3
1.2 Extensions using the ideas in the paper.	5
1.3 Organisation of paper.	6
1.4 Notation and background.	6
2. Space, time, and motion	9
2.1 Space models and body reference spaces.	10
2.2 Rigid transformations.	10
2.3 Rigid motions.	15

*Professor, DEPARTMENT OF MATHEMATICS AND STATISTICS, QUEEN’S UNIVERSITY, KINGSTON, ON K7L 3N6, CANADA

Email: andrew.lewis@queensu.ca, URL: <http://www.mast.queensu.ca/~andrew/>

Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada

3. Rigid bodies	19
3.1 Rigid bodies, mass, and centre of mass.	20
3.2 Inertia tensor.	22
3.3 Fibre semimetrics associated with a rigid body.	23
4. Rigid bodies with degenerate inertia tensors	24
4.1 Degeneracies of the inertia tensor and internal symmetry.	24
4.2 Reduced configuration space.	29
4.3 Reduced velocities.	31
4.4 Reduced kinetic energy metric.	36
5. Interconnected rigid body systems: kinematics	37
5.1 Configuration manifold.	37
5.2 Reduced configuration manifold.	38
5.3 Motion of interconnected rigid body systems.	40
5.4 Riemannian semimetrics and metrics for interconnected rigid body systems.	42
5.5 Velocity constraints for interconnected rigid body systems.	43
6. Interconnected rigid body systems: forces and torques	48
6.1 Forces and torques distributed on a body.	49
6.2 Primary torque-force.	54
6.3 Torque-force fields.	55
6.4 Work and power.	57
6.5 Virtual displacements, virtual power, and virtual work.	58
6.6 Interconnection torque-forces.	61
6.7 Constraint torque-forces.	61
7. Interconnected rigid body systems: dynamics	63
7.1 Momenta.	63
7.2 Newton–Euler equations for a single rigid body.	65
7.3 Newton–Euler equations for interconnected rigid body systems with velocity constraints.	68
7.4 Lagrange–d’Alembert Principle.	69
7.5 Kinetic energy.	70
8. Interconnected rigid body systems: Euler–Lagrange equations	72
8.1 Some constructions in variational calculus.	72
8.2 Affine connections, distributions, and submanifolds.	74
8.3 Mechanics on abstract Riemannian manifolds.	76
8.4 Euler–Lagrange equations for a single, free, unforced rigid body.	83
8.5 Lagrangian representations of torque-forces.	88
8.6 Equivalence of Newton–Euler and Euler–Lagrange equations.	90
References	94

1. Introduction

Our intention in this paper is to close the gap between the rich mathematical structures of geometric mechanics and the physical structures that give rise to them. The study of geometric mechanics typically begins with a mathematical framework, e.g., Riemannian geometry, symplectic or Poisson geometry, Dirac structures, etc., and develops this in various directions. Good examples of this are the books [Abraham and Marsden 1978, Bloch 2003, Libermann and Marle 1987]. With such treatments, the level of mathematical rigour and elegance is very appealing, and indeed is doubtless instrumental in the development of “geometric mechanics” becoming a field of mathematics in its own right. However, the entry point in the modelling process of such treatments is quite far along from what one might call “first principles.” On the other hand, the derivations of physical models from these first principles have been developed primarily in the engineering, applied mechanics, or physics literature. As such, they are typically presented in a language and style that is often difficult for a mathematician to understand [e.g., Goldstein 1951, Papastavridis 1999, Pars 1965, Whittaker 1937]. There have been various attempts to intermingle mathematical rigour with classical mechanics [e.g., Arnol’d 1989, Bullo and Lewis 2004, Murray, Li, and Sastry 1994, Spivak 2010]. Nonetheless, until now the project has not been fully undertaken to extend the principles of sound mathematics to all aspects of the basic physical modelling of mechanical systems.¹

1.1. Features of the presentation. We shall outline the existing work, while at the same time presenting what is novel about our approach.

1. As concerns essential mechanical ideas, the paper is completely self-contained. Mechanical elements are given mathematical definitions that we believe connect directly to physically meaningful constructions. That is to say, we do not skip the steps from the physical idea to its mathematical representation. In particular, our definitions of velocity constraints, and forces and torques give precise mathematical meaning to the way in which one models these physical notions in practice. We cannot point to a treatment in the existing literature where mathematical precision and a sound physical basis appear side-by-side in this way. Typically, for example, mathematical treatments of classical mechanics neglect any comprehensive treatment of force and torque.
2. Our physical axioms are uncontroversial for the subject, as indeed they are those that are derived from centuries of empirical verification and widespread use (outside the relativistic and quantum domains). Therefore, axioms differing in any material way from

¹In continuum mechanics, such an undertaking was made by Truesdell and Noll [1965], among others. Indeed, the continuum mechanics literature has been much better, in general, than the “analytical mechanics” literature about presenting its first principles in a mathematically rigorous manner. And, since analytical mechanics is, in principle, a special case of continuum mechanics, one might argue that what we do here has been done by Truesdell and Noll or in related work. We believe that works like that of Truesdell and Noll are fine pieces of work. We will, however, say two things. The first is that, while rigorous and comprehensive, the treatment of Truesdell and Noll—and indeed in much of the related written work on continuum field theories—is not extremely geometric. The path from a work such as that of Truesdell and Noll to a more geometric presentation, such as that of Marsden and Hughes [1983], requires more than just mere notational translation. And it is *geometric* mechanics that interests us here. Secondly, we will say that there are things that we do here that cannot be gleaned, directly or indirectly, from previous work. For example, our treatment of rigid bodies with degenerate inertia tensor cannot be found anywhere in the literature.

the ones we give, e.g., giving possibly different formulations of mechanics, will themselves be controversial. Moreover, we are not interested in philosophical discriminations between mathematically equivalent theories.

3. We make unreserved use of techniques from differential geometry, as these are natural in the physical modelling we undertake here, cf. [Lewis 2007]. Thus we extend the edifice of geometric mechanics all the way down to physical first principles. We believe this extension has been sorely missing from the field, and hope this will be a welcome diversion from the extension upwards of the superstructure of geometric mechanics.
4. We comprehensively develop the kinematics of rigid body motion in a general and abstract framework, using affine spaces. The generality and abstraction may be seen to make some things seem complicated, but we believe that clarity is attained by understanding exactly the mathematical representations of physical notions. There are several places where one can find a systematic presentation of space and time in Newtonian mechanics, including [Arnol'd 1989, §1.2] and [Artz 1981]. Much of this is dedicated to the mechanics of particles, which can be developed elegantly in a Galilean setting. The dynamics of rigid bodies in the Galilean setting is more difficult, and is explicated by Bhand and Lewis [2005]. We back off a little from a full presentation of Galilean mechanics since it is quite complicated, and really only becomes practical after the introduction of an observer, in which case one ends up with exactly our framework. (The difficulty, as explained in [Bhand and Lewis 2005], is that the Galilean group does not act effectively on spacetime, whereas the product of the Euclidean group with time translations does.)
5. We present a comprehensive treatment of rigid bodies. One of the novel contributions of the paper is to detail how one handles rigid bodies with degenerate inertia, i.e., particles and infinitely thin rods. The careful development of this requires some significant effort, and this has never been done. Moreover, in order to obtain a general theory, this procedure *must* be carried out, cf. Remark 4.10–2.
6. Nowhere in the paper will be found a utilisation of generalised coordinates or coordinates of any kind. In particular, we never have occasion to parameterise the rotation group.
7. The development of the Euler–Lagrange equations from Newtonian mechanics is a part of any moderately advanced treatment of classical mechanics. The arc of this development has become more or less standardised and has not really changed much in spirit from the original proof of Lagrange [1788]. In particular, published treatments of the connection between Newton and Euler–Lagrange are typically made in the context of particle mechanics, i.e., for mechanical systems comprised of interconnections of particles. The development of rigid body mechanics in this sort of framework can be done by replacing every rigid body with at most six particles, of appropriate mass and interconnected appropriately, cf. Example 4.4–3; however, this is unsatisfying. A derivation of the Euler–Lagrange equations from the Newton–Euler equations for a single unconstrained rigid body is undertaken by O’Reilly [2008, Chapter 10] using Euler angles.

What we do here is provide a development, free of coordinates, of the Euler–Lagrange equations for interconnections of rigid bodies (particles being special cases of rigid bodies in this framework), directly from the rigid body equations of Newton–Euler. We do this by a combination of variational methods and techniques from Riemannian geometry.

1.2. Extensions using the ideas in the paper. Apart from providing, we believe, some new insights into things that are classically (if only sometimes implicitly) well-understood, we also believe that the methodology of the paper can contribute to developments in mechanics that are not quite the classical ones. By making very clear mathematically the physical foundations of mechanics, we hope that one can develop mechanics in physical settings less restrictive than the usual one.

To this end, we list a few places where we think there may be progress to be made along the lines of what is presented in the paper.

1. In most developments of mechanics, including ours, one assumes a nice structure to the space of configurations, namely that it has a smooth manifold structure. However, this need not always be the case. One can have physical environments with boundaries, and different regimes in which the set of configurations undergoes an abrupt change. Such configuration spaces are “nonsmooth,” and our careful and precise construction of “physical configurations” (in Definition 5.2) should allow a physically natural—and mathematically precise—development of the analysis of nonsmooth configuration spaces.
2. The equations governing the dynamics of systems with velocity constraints are differential-algebraic equations. Such equations normally require further analysis to establish results concerning existence and uniqueness of solutions. Most work on the subject, including ours here, circumvents these problems by assuming that the constraint distribution has locally constant rank; in this case one can prove the usual sorts of existence and uniqueness results for physical motions, cf. Theorem 8.11. However, this is not always a physically valid assumption. For example, in “ N -trailer” problems, if two or more of the trailers have their wheels aligned in a 90° -turning position, the constraint force bundle will drop rank. This will lead to the constraint distribution itself not being smooth. As far as we know, this is a problem that is simply not understood, despite the fact that it is not uncommon in applications. A different sort of problem is when the constraint distribution is smooth but not locally constant rank. Such problems are considered by Cortés, de León, Martín de Diego, and Martínez [2001]. A key ingredient in dealing with such problems are physical assumptions that enable one to “close the loop” and arrive at consistent equations of motion. In general, the phrasing of physically meaningful and mathematically precise axioms governing a model is an important step in the modelling process. For example, our formulation of the Principle of Virtual Work in Definition 6.23 in a way that covers interconnection and constraint forces is just such a physically meaningful and mathematically precise axiom. This is, we believe, an improvement of the physically and mathematically nebulous form usually taken by the Principle of Virtual Work (but we refer to [Spivak 2010] for a detailed and reasonably clear, mathematically, description of topics surrounding the Principle of Virtual Work). Moreover, perhaps it can serve as a basis for the physical assumptions needed to correctly model systems where the constraint distribution is not smooth and of locally constant rank.
3. Appropriate physical principles can also be useful when dealing with generalising the sorts of forces one considers. For example, one can consider forces that are differential inclusions [Glocker 2001, Monteiro Marques 1993]. Our unusually careful development of forces and torques may be helpful in developing principles for handling general classes of forces.

4. The extension of our methodology to models leading to partial differential equations is appealing. As we mention in the footnote above, this is carried out for continuum mechanics by [Truesdell and Noll \[1965\]](#) and others. We believe that, nonetheless, a mathematical refinement of this work would serve to provide a useful and interesting foundation for the physical developments of these fields along geometric lines.

The first three of the preceding possible extensions based on our modelling all fall into the broad category of “nonsmooth mechanics.” This discipline was initiated in main part by work of Jean Jacques Moreau in the 1960’s, and his work continued into the 1980’s. There has been a lot of work in this area subsequently, and we refer to the books [[Brogliato 1999](#), [Kanno 2011](#)] and references contained therein for a survey of this work. However, there seems to be almost no work done on nonsmooth mechanics in the geometric setting. This seems to be a direction in which the field of geometric mechanics can be profitably extended.

1.3. Organisation of paper. The organisation of the paper is as follows. We describe our notion of spacetime, and transformations and motions of and through space in Section 2. In Section 3 we review what we mean by a rigid body, following, in spirit, the presentation of Bullo and Lewis [[2004](#), §4.2]. Special consideration is given in Section 4 to rigid bodies with degenerate inertia tensor. Here many of the new technical results in the paper are initiated. Single rigid bodies are generalised to interconnections of multiple rigid bodies in Section 5, where we define what we call an “interconnected rigid body system.” The emphasis in this section is on “kinematic” constructions, including velocity constraints. In Section 6 we provide a systematic presentation of forces on rigid bodies and in the context of interconnected systems with constraints. In Section 7 we present the Newton–Euler equations for the motion of an interconnection of rigid bodies subject to external forces and velocity constraints. The Euler–Lagrange equations, and a proof of their equivalence with the Newton–Euler equations, is the topic of Section 8. Included as an essential part of this equivalence are various versions of the “Lagrange–d’Alembert Principle.”

1.4. Notation and background. Our set theoretic conventions are standard, except that we denote set inclusion by “ \subseteq ,” with “ \subset ” standing for strict set inclusion. By id_X we denote the identity map on a set X . Sometimes, when the sets have cumbersome notation attached to them, we shall just denote the identity map by id , the set being understood from context. By 2^X we denote the set of subsets of a set X .

By \mathbb{Z} we denote the set of integers with $\mathbb{Z}_{>0}$ denoting the positive integers and $\mathbb{Z}_{\geq 0}$ denoting the nonnegative integers. By \mathbb{R} we denote the set of real numbers with $\mathbb{R}_{>0}$ denoting the set of positive real numbers.

If X is a topological space and if $A \subseteq X$, by $\text{int}(A)$ we denote the interior of A . If $A \subseteq B \subseteq X$, by $\text{int}_B(A)$ we denote the interior of A in B , i.e., the interior of A in the relative topology of B induced by X .

We suppose the reader to be familiar with standard linear algebra. Our vector spaces will typically be over the field of real numbers. By $\text{Hom}_{\mathbb{R}}(\mathbf{U}; \mathbf{V})$ we denote the space of \mathbb{R} -linear maps between \mathbb{R} -vector spaces \mathbf{U} and \mathbf{V} . We abbreviate $\mathbf{V}^* = \text{Hom}_{\mathbb{R}}(\mathbf{V}; \mathbb{R})$ and $\text{End}_{\mathbb{R}}(\mathbf{V}) = \text{Hom}_{\mathbb{R}}(\mathbf{V}; \mathbf{V})$. If $\alpha \in \mathbf{V}^*$ and $v \in \mathbf{V}$, we will use either $\langle \alpha; v \rangle$ or $\alpha(v)$ to denote

the evaluation of α on v . If $S \subseteq \mathbf{V}$, then we denote by

$$\text{ann}(S) = \{\alpha \in \mathbf{V}^* \mid \alpha(v) = 0 \text{ for all } v \in S\}$$

the *annihilator* of S . If $S \subseteq \mathbf{V}$, by $\text{span}(S)$ we denote the subspace generated by S . For $A \in \text{Hom}_{\mathbb{R}}(\mathbf{U}; \mathbf{V})$, we denote the dual of A by $A^*: \mathbf{V}^* \rightarrow \mathbf{U}^*$, and recall that it is defined to satisfy

$$\langle A^*(\beta); u \rangle = \langle \beta; A(u) \rangle, \quad u \in \mathbf{U}, \beta \in \mathbf{V}^*.$$

By $\bigwedge^k(\mathbf{V}^*)$ we denote the space of alternating k -forms on \mathbf{V} , which we think of as being the alternating multilinear mappings from \mathbf{V}^k into \mathbb{R} . That is, we identify $\bigwedge^k(\mathbf{V}^*)$ with the mappings

$$A: \underbrace{\mathbf{V} \times \cdots \times \mathbf{V}}_{k \text{ times}} \rightarrow \mathbb{R}$$

that are linear in each entry and which satisfy

$$A(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \text{sign}(\sigma)A(v_1, \dots, v_k)$$

for every permutation σ of the set $(1, \dots, k)$, where $\text{sign}(\sigma)$ denotes the parity of the permutation.

We shall rely heavily on constructions special to vector spaces with inner products and orientations, and here outline our notation for these. We let \mathbf{V} be an n -dimensional \mathbb{R} -vector space. An *orientation* of \mathbf{V} is an equivalence class in $\bigwedge^n(\mathbf{V}^*) \setminus \{0\}$, where nonzero n -forms are equivalent if they agree up to a positive scalar multiple (keeping in mind that $\dim_{\mathbb{R}}(\bigwedge^n(\mathbf{V}^*)) = 1$). We shall designate the choice of orientation by a single nonzero n -form θ , taking for granted that it is understood that it is the equivalence class of θ that we have in mind. A basis (e_1, \dots, e_n) of \mathbf{V} is *positively-oriented* if $\theta(e_1, \dots, e_n) \in \mathbb{R}_{>0}$. An invertible linear map $A \in \text{End}_{\mathbb{R}}(\mathbf{V})$ is *orientation-preserving* if $A^*\theta$ defines the same orientation as θ , where

$$A^*\theta(v_1, \dots, v_n) = \theta(A(v_1), \dots, A(v_n)).$$

Now let \mathbf{V} be a finite-dimensional \mathbb{R} -vector space with inner product g and orientation θ . Denote

$$\|v\| = \sqrt{g(v, v)}, \quad v \in \mathbf{V}.$$

By $\text{SO}(\mathbf{V}, g, \theta)$ we denote the orientation-preserving linear isometries of \mathbf{V} . Thus $A \in \text{SO}(\mathbf{V}, g, \theta)$ if A is orientation-preserving and additionally satisfies

$$g(A(v_1), A(v_2)) = g(v_1, v_2), \quad v_1, v_2 \in \mathbf{V}.$$

We refer to [Berger 1987, §8.2, 8.11] for basic properties of $\text{SO}(\mathbf{V}, g, \theta)$. We denote by $\mathfrak{so}(\mathbf{V}, g, \theta)$ the subspace of $\text{End}_{\mathbb{R}}(\mathbf{V})$ comprised of g -skew-symmetric linear mappings, i.e., those satisfying

$$g(A(v_1), v_2) = -g(v_1, A(v_2)), \quad v_1, v_2 \in \mathbf{V}.$$

We denote by $g^b: \mathbf{V} \rightarrow \mathbf{V}^*$ the natural isomorphism defined by

$$\langle g^b(u); v \rangle = g(v, u), \quad u, v \in \mathbf{V}.$$

The inverse of g^\flat we denote by g^\sharp . If $U \subseteq V$ is a subspace, then U^{\perp_g} denotes the g -orthogonal complement of U . We denote by g^{-1} the inner product on V^* induced by g :

$$g^{-1}(\alpha, \beta) = g(g^\sharp(\alpha), g^\sharp(\beta)), \quad \alpha, \beta \in V^*.$$

Our principle interest is in 3-dimensional oriented vector spaces with an inner product g . We aggregate the data in a triple (V, g, θ) . In this case, we define a bilinear mapping $(v_1, v_2) \mapsto v_1 \times v_2 \in V$ on $V \times V$ as follows. First of all, for $u, v \in V$, define $n_{u,v} \in V$ by asking that $n \in (\text{span}(u, v))^{\perp_g}$ and that $(u, v, n_{u,v})$ be a positively-oriented orthogonal basis if u and v are not collinear, and $n_{u,v} = 0$ if u and v are collinear. Then define $\theta_{u,v} = \cos^{-1}(\frac{g(u,v)}{\|u\|\|v\|})$ and

$$u \times v = \|u\|\|v\| \sin(\theta_{u,v})n_{u,v}.$$

We refer to [Berger 1987, §8.11] for details and generalisations. We can directly verify the identity

$$g(u \times v, w) = g(w \times u, v), \quad u, v, w \in V, \quad (1.1)$$

that we shall frequently use. We define an injective homomorphism $v \mapsto \hat{v}$ from V into $\text{End}_{\mathbb{R}}(V)$ by requiring that $\hat{v}(u) = v \times u$ for $u \in V$. One can use (1.1) to prove that $\hat{v} \in \mathfrak{so}(V, g, \theta)$. Moreover, a dimension count shows that $\hat{\cdot}$ is surjective onto $\mathfrak{so}(V, g, \theta)$, and so is an isomorphism onto this latter space. The inverse map from $\mathfrak{so}(V, g, \theta)$ to V we denote by $A \mapsto \check{A}$.

While we shall mostly work with abstract vector spaces, when we work with \mathbb{R}^n we denote by $\mathbf{x} = (x_1, \dots, x_n)$ a typical point in \mathbb{R}^n . By $\mathbb{R}^{m \times n}$ we denote the set of $m \times n$ matrices with entries in \mathbb{R} , which we think of as being members of $\text{Hom}_{\mathbb{R}}(\mathbb{R}^n; \mathbb{R}^m)$. The standard basis for \mathbb{R}^n we denote by (e_1, \dots, e_n) .

We also work with affine spaces, following [Berger 1987, Chapter 2]. We recall that an **affine space** A modelled on a vector space V is defined by an effective, transitive action

$$A \times V \ni (x, v) \mapsto x + v \in A$$

of the Abelian group V on A . (Note that this is the *definition* of “+”!) Note that effectiveness and transitivity of the action ensures that, given $x_1, x_2 \in A$, there exists a unique $v \in V$ such that $x_2 = x_1 + v$. Thus we denote this unique v by $x_2 - x_1$. That is, we can (1) add elements of V to elements of A , which we think of as “translating” the element of A by the element of V and (2) subtract elements of A , which we think of as returning the amount we must translate to get from one point to the other. If the model vector space V has an inner product g , we define a metric on A by

$$d(x_1, x_2) = \sqrt{g(x_2 - x_1, x_2 - x_1)}. \quad (1.2)$$

Other constructions concerning affine spaces will be introduced in the text and notation provided for these as required. If $S \subseteq A$ is a subset of an affine space, by

$$\text{conv}(S) = \{x_1 + s(x_2 - x_1) \mid s \in [0, 1], x_1, x_2 \in S\}$$

we denote the convex hull of S , i.e., the union of all line segments between points of S . By

$$\text{aff}(S) = \{x_1 + s(x_2 - x_1) \mid s \in \mathbb{R}, x_1, x_2 \in S\}$$

we denote the affine hull of S , i.e., the union of all lines through points in S .

It is convenient in two places in our presentation to use measures, one when discussing mass distributions for rigid bodies, and the other for discussing forces and torques. We refer to [Cohn 2013] as a basic reference for measure theory. By $\text{supp}(\mu)$ we denote the support of a measure μ . One non-basic element of measure theory we shall use is that of a vector measure. Let $(\mathcal{M}, \mathcal{A})$ be a measurable space with \mathbb{T} a topological vector space. A *vector measure* with values in \mathbb{T} is a mapping $\mu: \mathcal{A} \rightarrow \mathbb{T}$ such that, if $(A_j)_{j \in \mathbb{Z}_{>0}}$ is a sequence of pairwise disjoint measurable sets, then

$$\mu(\cup_{j \in \mathbb{Z}_{>0}} A_j) = \sum_{j=1}^{\infty} \mu(A_j),$$

the sum converging in the topology of \mathbb{T} . We note that, if λ is an element of the continuous dual of \mathbb{T} , then we define a signed measure μ_λ on $(\mathcal{M}, \mathcal{A})$ by $\mu_\lambda(A) = \langle \lambda; \mu(A) \rangle$.

We shall assume the reader is familiar with basic differential geometry, and use [Abraham, Marsden, and Ratiu 1988] as a reference whose notational conventions we mainly adopt. We shall not often work with general manifolds, but when we do we shall suppose them to be smooth, Hausdorff, and second countable. By $\pi_{\text{TM}}: \text{TM} \rightarrow \text{M}$ we denote the tangent bundle of a manifold M and by $\pi_{\text{T}^*\text{M}}: \text{T}^*\text{M} \rightarrow \text{M}$ the cotangent bundle. The fibres of these bundles at $x \in \text{M}$ we denote by T_xM and T_x^*M , respectively. We shall sometimes denote the zero vector in a tangent or cotangent space at x by 0_x . For a subset $A \subseteq \text{M}$, we denote

$$\text{TM}|A = \{v_x \in \text{TM} \mid x \in A\}.$$

If $\Phi: \text{M} \rightarrow \text{N}$ is a differentiable mapping of manifolds, by $T\Phi: \text{TM} \rightarrow \text{TN}$ we denote the derivative. We also denote $T_x\Phi = T\Phi|_{\text{T}_x\text{M}}$. If $I \subseteq \mathbb{R}$ is an interval and if $\gamma: I \rightarrow \text{M}$ is a differentiable curve, then we define $\gamma': I \rightarrow \text{TM}$ by $\gamma'(t) = T_t\gamma \cdot 1$. If X is a vector field and f is a function, by $\mathcal{L}_X f$ we denote the Lie derivative of f with respect to X .

We shall make some reference to elementary Riemannian and affine differential geometry and we refer to [Bullo and Lewis 2004, §3.8] for a treatment at the level we shall use here, and for notation. For a Riemannian metric \mathbf{G} on a manifold M , we denote the Levi-Civita affine connection by $\overset{\mathbf{G}}{\nabla}$. An arbitrary affine connection we denote by ∇ . Thus $\nabla_X Y$ denotes the covariant derivative of Y with respect to X . For a curve $\gamma: I \rightarrow \text{M}$ and a vector field $Y: I \rightarrow \text{TM}$ along γ —i.e., $Y(t) \in \text{T}_{\gamma(t)}\text{M}$ —we denote by $\nabla_{\gamma'(t)} Y(t)$ the covariant derivative of Y along γ . Thus, in particular, geodesics are curves γ that satisfy $\nabla_{\gamma'(t)} \gamma'(t) = 0$.

Acknowledgements. The author wishes to thank James Forbes and Anton de Ruiter for reminding him that this work was (1) not done and (2) interesting. Jonny Briggs went through much of the manuscript carefully and found numerous errors, although those that remain are fully the responsibility of the author.

2. Space, time, and motion

We begin the technical presentation by carefully formulating our notions of space, time, motion, and velocity. We do not do this in the most general possible setting, which would be that of a Galilean spacetime. Rigid body dynamics in Galilean spacetimes is formulated

by [Bhand and Lewis \[2005\]](#). Here we suppose that spacetime is the product of the temporal axis \mathbb{T} and a spatial three-dimensional affine space. This means that we have established a notion of “stationary,” but that we have established neither a spatial nor a temporal origin.

2.1. Space models and body reference spaces. In order to prevent a proliferation of \mathbb{R}^3 's, all having different physical meanings, we shall use an abstract model for space.

2.1 Definition: (Newtonian space model) A *Newtonian space model* is a quadruple $\mathcal{S} = (\mathbb{S}, \mathbb{V}, g, \theta)$, where (\mathbb{V}, g) is a three-dimensional inner product space, θ is a nonzero member of $\bigwedge^3(\mathbb{V}^*)$ (defining an orientation on \mathbb{V}), and \mathbb{S} is an affine space modelled on \mathbb{V} . •

The affine space \mathbb{S} serves as our model for points in space.

We can make the abstraction of the preceding space model more concrete by use of the following device.

2.2 Definition: (Spatial frame) A *spatial frame* for a Newtonian space model $\mathcal{S} = (\mathbb{S}, \mathbb{V}, g, \theta)$ is a pair $(O_{\text{spatial}}, (s_1, s_2, s_3))$, where $O_{\text{spatial}} \in \mathbb{S}$ and (s_1, s_2, s_3) is a positively-oriented orthonormal basis for \mathbb{V} . •

Given a spatial frame, for each $x \in \mathbb{S}$ we can write

$$x = O_{\text{spatial}} + x^1 s_1 + x^2 s_2 + x^3 s_3$$

for some unique $x^1, x^2, x^3 \in \mathbb{R}$. This, then, establishes a bijection between \mathbb{S} and \mathbb{R}^3 that can be useful in making some of our abstract constructions concrete. We shall, however, not explicitly pursue making this translation from abstract to concrete in this paper.

In the preceding discussion, we described our model of physical space. We will posit a similar model for the reference space in which a rigid body resides.

2.3 Definition: (Body reference space) A *body reference space* is a quadruple $\mathcal{B} = (\mathbb{B}, \mathbb{U}, G, \Theta)$, where (\mathbb{U}, G) is a three-dimensional inner product space, Θ is a nonzero member of $\bigwedge^3(\mathbb{U}^*)$, and \mathbb{B} is an affine space modelled on \mathbb{U} . •

We also have the corresponding notion of a reference frame for a body reference space.

2.4 Definition: (Body frame) A *body frame* for a body reference space $\mathcal{B} = (\mathbb{B}, \mathbb{U}, G, \Theta)$ is a pair $(O_{\text{body}}, (b_1, b_2, b_3))$, where $O_{\text{body}} \in \mathbb{B}$ and (b_1, b_2, b_3) is a positively-oriented orthonormal basis for \mathbb{U} . •

2.2. Rigid transformations. We now present the manner in which we will describe the motion of a rigid body. We suppose that we have a Newtonian space model $\mathcal{S} = (\mathbb{S}, \mathbb{V}, g, \theta)$ and a body reference space $\mathcal{B} = (\mathbb{B}, \mathbb{U}, G, \Theta)$. The notion of a rigid motion is described by relating these spaces, so the body is not actually required for this; it merely “goes along for the ride” after the fact. Thus we only introduce bodies in Section 3.

2.5 Definition: (Rigid transformation) For a Newtonian space model $\mathcal{S} = (\mathbb{S}, \mathbb{V}, g, \theta)$ and a body reference space $\mathcal{B} = (\mathbb{B}, \mathbb{U}, G, \Theta)$, a *rigid transformation* of \mathcal{B} in \mathcal{S} is a map $\Phi: \mathbb{B} \rightarrow \mathbb{S}$ having the following properties:

- (i) Φ is an affine map, i.e., there exists $R_\Phi \in \text{Hom}_{\mathbb{R}}(\mathbb{U}; \mathbb{V})$ such that

$$\Phi(X) = \Phi(X_0) + R_\Phi(X - X_0), \quad X \in \mathbb{B}, \quad (2.1)$$

for any $X_0 \in \mathbb{B}$;

- (ii) $R_\Phi^* g = G$;
- (iii) $R_\Phi^* \theta = \alpha \Theta$ for some $\alpha \in \mathbb{R}_{>0}$.

By $\text{Rgd}(\mathcal{B}; \mathcal{S})$ we denote the set of rigid transformations of \mathcal{B} in \mathcal{S} . •

We note that, for Φ to be affine, it suffices to check that (2.1) holds for *some* $X_0 \in \mathbf{B}$ [Berger 1987, Proposition 2.3.1]. We also note that this uniquely defines $R_\Phi \in \text{Hom}_{\mathbb{R}}(\mathbf{U}; \mathbf{V})$ [Berger 1987, Proposition 2.3.1]. The last two conditions mean that R_Φ is an orientation-preserving isometry. With this mind, let us denote by $\text{Isom}^+(\mathcal{B}, \mathcal{S})$ the set of linear orientation-preserving isometries from \mathbf{U} to \mathbf{V} . We will also have occasion to refer to the groups of orientation preserving isometries of \mathbf{B} and \mathbf{S} , which we denote by $\text{Rgd}(\mathcal{B})$ and $\text{Rgd}(\mathcal{S})$, respectively. Note that these are indeed groups (with the operation of composition), but that $\text{Rgd}(\mathcal{B}; \mathcal{S})$ is not naturally a group. As previously, $\text{SO}(\mathbf{U}, G, \Theta)$ and $\text{SO}(\mathbf{V}, g, \theta)$ denote the groups of orientation-preserving linear isometries of \mathbf{U} and \mathbf{V} , respectively.

To understand the structure of the space of rigid transformations, the following result provides a useful insight.

2.6 Proposition: (Group actions on the space of rigid transformations) *Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, g, \theta)$ be a Newtonian space model and let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, G, \Theta)$ be a body reference space. Then the mappings*

$$\begin{aligned} \mathcal{A}_{\text{spatial}}: \text{Rgd}(\mathcal{S}) \times \text{Rgd}(\mathcal{B}; \mathcal{S}) &\rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S}) \\ (\Psi, \Phi) &\mapsto \Psi \circ \Phi \end{aligned}$$

and

$$\begin{aligned} \mathcal{A}_{\text{body}}: \text{Rgd}(\mathcal{B}; \mathcal{S}) \times \text{Rgd}(\mathcal{B}) &\rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S}) \\ (\Phi, \Psi) &\mapsto \Phi \circ \Psi \end{aligned}$$

define left- and right-actions of $\text{Rgd}(\mathcal{S})$ and $\text{Rgd}(\mathcal{B})$, respectively, on $\text{Rgd}(\mathcal{B}; \mathcal{S})$. Moreover, both actions are free and transitive.

Proof: That $\mathcal{A}_{\text{spatial}}$ and $\mathcal{A}_{\text{body}}$ are actions is easily verified. Let us verify the freeness and transitivity of $\mathcal{A}_{\text{spatial}}$, the similar assertions for $\mathcal{A}_{\text{body}}$ following in a similar manner. Let $\Psi \in \text{Rgd}(\mathcal{S})$ have the property that $\Psi \circ \Phi = \Phi$ for every $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$. Let $x \in \mathbf{S}$ and note that

$$\Psi(x) = \Psi \circ \Phi \circ \Phi^{-1}(x) = \Phi \circ \Phi^{-1}(x) = x,$$

giving $\Psi = \text{id}_{\mathbf{S}}$. This proves freeness of $\mathcal{A}_{\text{spatial}}$. Next let $\Phi_1, \Phi_2 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ and note that $\Phi_2 \circ \Phi_1^{-1} \in \text{Rgd}(\mathcal{S})$ since compositions of orientation-preserving isometries are orientation-preserving isometries. Moreover,

$$\mathcal{A}_{\text{spatial}}(\Phi_2 \circ \Phi_1^{-1}, \Phi_1) = \Phi_2,$$

giving transitivity of $\mathcal{A}_{\text{spatial}}$. ■

We now give a series of lemmata that give further structure to the set of rigid transformations. We begin by showing that the set of rigid transformations can be given the structure of a group by fixing a reference rigid transformation with respect to which others are measured.

2.7 Lemma: (Rigid transformations with respect to a reference transformation)

Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model and let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space. Let $\Phi_0 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$. Then the binary operation

$$(\Phi, \Psi) \mapsto \Phi * \Psi \triangleq \Phi \circ \Phi_0^{-1} \circ \Psi$$

makes $\text{Rgd}(\mathcal{B}; \mathcal{S})$ into a group. Moreover, the mapping $\Phi \mapsto \Phi \circ \Phi_0^{-1}$ is an isomorphism of this group with $\text{Rgd}(\mathcal{S})$ and the mapping $\Phi \mapsto \Phi_0^{-1} \circ \Phi$ is an isomorphism of this group with $\text{Rgd}(\mathcal{B})$.

Proof: The group operation is associative:

$$\begin{aligned} \Phi * (\Psi * \Gamma) &= \Phi \circ \Phi_0^{-1} \circ (\Psi \circ \Phi_0^{-1} \circ \Gamma) \\ &= (\Phi \circ \Phi_0^{-1} \circ \Psi) \circ \Phi_0^{-1} \circ \Gamma \\ &= (\Phi * \Psi) * \Gamma. \end{aligned}$$

The identity element is Φ_0 :

$$\Phi_0 * \Phi = \Phi_0 \circ \Phi_0^{-1} \circ \Phi = \Phi, \quad \Phi * \Phi_0 = \Phi \circ \Phi_0^{-1} \circ \Phi_0 = \Phi.$$

The inverse of Φ is $\Phi_0 \circ \Phi^{-1} \circ \Phi_0$:

$$\begin{aligned} \Phi * (\Phi_0 \circ \Phi^{-1} \circ \Phi_0) &= \Phi \circ \Phi_0^{-1} \circ \Phi_0 \circ \Phi^{-1} \circ \Phi_0 = \Phi_0, \\ (\Phi_0 \circ \Phi^{-1} \circ \Phi_0) * \Phi &= \Phi_0 \circ \Phi^{-1} \circ \Phi_0 \circ \Phi_0^{-1} \circ \Phi = \Phi_0, \end{aligned}$$

giving the first part of the result.

For the second assertion, we note that compositions of isometries (resp. orientation-preserving maps) are again isometries (resp. orientation-preserving maps). \blacksquare

The group described in the preceding lemma we denote by $\text{Rgd}_{\Phi_0}(\mathcal{B}; \mathcal{S})$. The lemma establishes isomorphisms of this group with *both* $\text{Rgd}(\mathcal{S})$ and $\text{Rgd}(\mathcal{B})$. In this paper we shall focus on the former representation, although it is also possible to work with the latter.

In order to understand better the structure of a rigid transformation, we have the following result.

2.8 Lemma: (Rigid transformations with respect to origins) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model and let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space. Let $x_0 \in \mathbb{S}$ and let $X_0 \in \mathbb{B}$. If $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, then there exist $r_\Phi \in \mathbb{V}$ and $R_\Phi \in \text{Isom}^+(\mathbb{B}; \mathcal{S})$ such that

$$\Phi(X) = x_0 + (r_\Phi + R_\Phi(X - X_0)), \quad X \in \mathbb{B}. \quad (2.2)$$

Moreover, R_Φ is uniquely determined by Φ and does not depend on x_0 or X_0 .

Proof: First note that, for any $r \in \mathbb{V}$ and $R \in \text{Hom}_{\mathbb{R}}(\mathbb{U}; \mathbb{V})$, the mapping

$$\mathbb{B} \ni X \mapsto x_0 + (r + R(X - X_0)) \in \mathbb{S}$$

is affine. Let us take $r_\Phi = \Phi(X_0) - x_0$ and R_Φ to be the linear mapping in Definition 2.5, and define

$$\Psi(X) = x_0 + (r_\Phi + R_\Phi(X - X_0)).$$

Note that $\Psi(X_0) = x_0 + r_\Phi = \Phi(X_0)$ and so

$$\Psi(X) = \Phi(X_0) + R_\Phi(X - X_0), \quad X \in \mathcal{B},$$

giving $\Psi = \Phi$.

That R_Φ does not depend on x_0 or X_0 follows since R_Φ is the linear part of Φ . \blacksquare

The result says that, upon a choice of body and spatial origin, one can think of a rigid transformation as a rotation followed by a translation.

Whenever we specify a body origin, we shall use this to assign $R_\Phi \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$ to $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, and we shall do so without explicitly mentioning that this is done via the preceding lemma. If we additionally choose a spatial origin x_0 , we shall assign $r_\Phi \in \mathcal{V}$ to $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, again without making explicit mention that this is done by the lemma.

The preceding two lemmata can be combined to give the following description of the space of rigid transformations.

2.9 Lemma: (A concrete group representation for rigid transformations) *Let $\mathcal{S} = (\mathcal{S}, \mathcal{V}, \mathfrak{g}, \theta)$ be a Newtonian space model and let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathfrak{G}, \Theta)$ be a body reference space. Let $x_0 \in \mathcal{S}$, let $X_0 \in \mathcal{B}$, let $\Phi_0 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, and define $(R_0, r_0) \in \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathcal{V}$ by*

$$\Phi_0(X) = x_0 + (r_0 + R_0(X - X_0)).$$

Then the mapping $\Phi \mapsto (R_\Phi \circ R_0^{-1}, r_\Phi - R_\Phi \circ R_0^{-1}(r_0))$ is a group isomorphism from $\text{Rgd}_{\Phi_0}(\mathcal{B}; \mathcal{S})$ to $\text{SO}(\mathcal{V}, \mathfrak{g}, \theta) \times \mathcal{V}$, if the latter has the group operation

$$(R, r) \cdot (S, s) = (R \circ S, r + R(s)). \quad (2.3)$$

Proof: We have

$$\Phi_0^{-1}(x) = X_0 + R_0^{-1}(x - x_0 - r_0)$$

and so

$$\Phi \circ \Phi_0^{-1}(x) = x_0 + (r_\Phi - R_\Phi \circ R_0^{-1}(r_0) + R_\Phi \circ R_0^{-1}(x - x_0)).$$

From this one directly computes that

$$\begin{aligned} (\Phi \circ \Phi_0^{-1}) \circ (\Psi \circ \Phi_0^{-1})(x) &= x_0 + (r_\Phi - R_\Phi \circ R_0^{-1}(r_0) \\ &\quad + R_\Phi \circ R_0^{-1}(r_\Psi - R_\Psi \circ R_0^{-1}(r_0)) + (R_\Phi \circ R_0^{-1}) \circ (R_\Psi \circ R_0^{-1})(x - x_0)), \end{aligned}$$

from which the lemma follows. \blacksquare

The preceding lemma says that, upon a choice of spatial and body origin, and of a reference configuration, the set of rigid transformations becomes the well-known semidirect product of rotations and translations.

Let us make the representation of rigid transformations even more concrete by the use of spatial and body frames.

2.10 Lemma: (Rigid transformations with respect to frames) Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model with spatial frame $(O_{\text{spatial}}, (s_1, s_2, s_3))$, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space with body frame $(O_{\text{body}}, (b_1, b_2, b_3))$, and let $\Phi_0 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ be defined by $r_0 = 0$ and $R_0(b_j) = s_j$, $j \in \{1, 2, 3\}$, where r_0 and R_0 are as in Lemma 2.9. Let r_Φ and R_Φ be as in Lemma 2.8, taking $x_0 = O_{\text{spatial}}$ and $X_0 = O_{\text{body}}$. Let $\mathbf{r}_\Phi \in \mathbb{R}^3$ be the vector of components of r_Φ with respect to the basis (s_1, s_2, s_3) and let $\mathbf{R}_\Phi \in \text{SO}(3)$ be the matrix of R_Φ with respect to the bases (b_1, b_2, b_3) and (s_1, s_2, s_3) . Then the following statements hold:

- (i) $\Phi(O_{\text{body}}) - O_{\text{spatial}} = r^1 s_1 + r^2 s_2 + r^3 s_3$;
- (ii) the (i, j) th component of \mathbf{R}_Φ is the i th component of $R_\Phi(b_j)$ relative to the basis (s_1, s_2, s_3) ;
- (iii) the components of the vector $\Phi(X) - O_{\text{spatial}} \in \mathbf{V}$ with respect to the basis (s_1, s_2, s_3) are the components of vector $\mathbf{r}_\Phi + \mathbf{R}_\Phi \mathbf{X}$, where \mathbf{X} are the components of the vector $X - O_{\text{body}}$ with respect to the basis (b_1, b_2, b_3) .

Proof: (i) In the proof of Lemma 2.8 we showed that $r_\Phi = \Phi(O_{\text{body}}) - O_{\text{spatial}}$, and this immediately gives this part of the lemma.

(ii) Let R_j^i be the (i, j) th component of \mathbf{R}_Φ . By definition we have

$$R_\Phi(b_j) = R_j^1 s_1 + R_j^2 s_2 + R_j^3 s_3, \quad j \in \{1, 2, 3\},$$

and this part of the result follows directly from this.

(iii) As in Lemma 2.8 we have

$$\Phi(X) - O_{\text{spatial}} = r_\Phi + R_\Phi(X - X_0),$$

and the conclusion follows by representing this equation in the bases (b_1, b_2, b_3) for \mathbf{U} and (s_1, s_2, s_3) for \mathbf{V} . ■

The situation of the lemma is depicted in Figure 1. Note that the lemma establishes

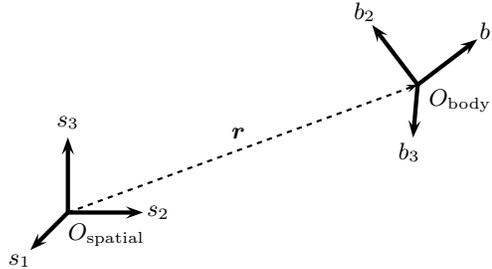


Figure 1. A rigid transformation with spatial and body frames

a correspondence between $\text{Rgd}(\mathcal{B}; \mathcal{S})$ and $\text{SO}(3) \times \mathbb{R}^3$ by a choice of spatial and body frames.

We now have at hand three groups connected with the set of rigid transformations:

1. $\text{Rgd}_{\Phi_0}(\mathcal{B}; \mathcal{S})$ with the group operation of Lemma 2.7 that depends on a choice of $\Phi_0 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$;
2. $\text{SO}(\mathbb{V}, \mathfrak{g}, \theta) \times \mathbb{V}$ with the group operation

$$(R_1, v_1) \cdot (R_2, v_2) = (R_1 \circ R_2, v_1 + R_1(v_2));$$

3. $\text{SO}(3) \times \mathbb{R}^3$ with the group operation

$$(\mathbf{R}_1, \mathbf{v}_1) \cdot (\mathbf{R}_2, \mathbf{v}_2) = (\mathbf{R}_1 \mathbf{R}_2, \mathbf{v}_1 + \mathbf{R}_1 \mathbf{v}_2).$$

The preceding lemmata establish identifications of each of these groups with $\text{Rgd}(\mathcal{B}; \mathcal{S})$ by various choices, namely

1. a choice of $\Phi_0 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$,
2. a choice of $x_0 \in \mathbb{S}$, $X_0 \in \mathbb{B}$, and $\Phi_0 \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, or
3. the data from 2 plus a choice $(O_{\text{spatial}}, (s_1, s_2, s_3))$ of spatial frame and $(O_{\text{body}}, (b_1, b_2, b_3))$ of body frame, respectively.

The final, most concrete, representation of a rigid transformation is that which is most commonly used [e.g., Bullo and Lewis 2004, Murray, Li, and Sastry 1994]. However, there is substantial clarification in working with the more abstract representations, especially when we come to talk about degenerate rigid bodies in Section 4.1 below. Indeed, we shall not for the remainder of the paper make any essential reference to spatial or body frames.

2.3. Rigid motions. If we put things in motion, this requires time.

2.11 Definition: (Time) A *time axis* is an affine space \mathbb{T} modelled on \mathbb{R} .² A *time interval* is a subset $\mathbb{T}' \subseteq \mathbb{T}$ of the form

$$\mathbb{T}' = \{t_0 + t \mid t \in I\}$$

for some $t_0 \in \mathbb{T}$ and for some interval $I \subseteq \mathbb{R}$. •

Suppose that we have an affine space \mathbb{A} modelled on a finite-dimensional \mathbb{R} -vector space \mathbb{V} and a curve $\phi: \mathbb{T}' \rightarrow \mathbb{A}$ from a time interval $\mathbb{T}' \subseteq \mathbb{T}$. The curve is *differentiable* at $t_0 \in \mathbb{T}'$ if the limit

$$\lim_{t \rightarrow t_0} \frac{\phi(t) - \phi(t_0)}{t - t_0}$$

exists. Note that the limit, when it exists, is necessarily in \mathbb{V} , and we denote this limit by $\dot{\phi}(t_0)$. Note that $\dot{\phi}$ and ϕ' are not the same thing, and the reader will want to keep this in mind at some points. If $\dot{\phi}: \mathbb{T}' \rightarrow \mathbb{V}$ is differentiable, then we denote

$$\ddot{\phi}(t_0) = \lim_{t \rightarrow t_0} \frac{\dot{\phi}(t) - \dot{\phi}(t_0)}{t - t_0},$$

when this limit exists. Again this limit is in \mathbb{V} when it exists, and when it exists we say that ϕ is *twice differentiable* at t_0 .

Next we consider the situation when a rigid body is in motion.

²One may replace the model vector space \mathbb{R} with a more general object, namely a one-dimensional oriented inner product space. The orientation, in this case, serves the purpose of distinguishing vectors that point “forward” in time. However, one gains literally nothing by doing this. Indeed, given a one-dimensional oriented inner product space, there exists a *unique* orientation preserving isometry with \mathbb{R} , equipped with its standard inner product and equipped with the standard orientation “dt.”

2.12 Definition: (Rigid motion) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathfrak{G}, \Theta)$ be a body reference space, and let \mathbb{T} be a time axis. A *rigid motion* of \mathcal{B} in \mathcal{S} is a curve $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ defined on a time interval $\mathbb{T}' \subseteq \mathbb{T}$. \bullet

As we saw in the preceding section, by a choice of spatial and body frame, we establish an identification $\text{Rgd}(\mathcal{B}; \mathcal{S}) \simeq \text{SO}(3) \times \mathbb{R}^3$. Using this identification, we can assign regularity to a rigid motion ϕ by considering its regularity when represented by a curve ϕ in $\text{SO}(3) \times \mathbb{R}^3$. Since two representations of $\text{Rgd}(\mathcal{B}; \mathcal{S})$ by $\text{SO}(3) \times \mathbb{R}^3$ are related by a smooth diffeomorphism (as is easily seen), any notion of regularity up to infinitely differentiable³ will be independent of choices of spatial and body frames. Thus we shall freely speak of things like “differentiable rigid motions.” The upshot of the preceding, of course, is that $\text{Rgd}(\mathcal{B}; \mathcal{S})$ is a smooth manifold that is diffeomorphic to $\text{SO}(3) \times \mathbb{R}^3$ via any choice of spatial and body frames.

Let us make some observations about velocities associated with rigid motions.

2.13 Lemma: (Velocities associated with a rigid motion) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathfrak{G}, \Theta)$ be a body reference space, and let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. For a differentiable rigid motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$, and for $x_0 \in \mathbb{S}$ and $X_0 \in \mathbb{B}$, let $R_\phi: \mathbb{T}' \rightarrow \text{Isom}^+(\mathcal{B}; \mathcal{S})$ and $r_\phi: \mathbb{T}' \rightarrow \mathbb{V}$ satisfy

$$\phi(t)(X) = x_0 + (r_\phi(t) + R_\phi(t)(X - X_0)), \quad X \in \mathbb{B},$$

as in Lemma 2.8. Then the following statements hold:

- (i) $R_\phi^T(t) \circ \dot{R}_\phi(t) \in \mathfrak{so}(\mathbb{U}; \mathfrak{G}, \Theta)$ and $\dot{R}_\phi(t) \circ R_\phi^T(t) \in \mathfrak{so}(\mathbb{V}, \mathfrak{g}, \theta)$;
- (ii) \dot{r}_ϕ depends only on ϕ and X_0 , and not on x_0 .

Proof: (i) Since $R_\phi(t) \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$ for each $t \in \mathbb{T}'$, we have

$$\mathfrak{g}(R_\phi(t)(u_1), R_\phi(t)(u_2)) = \mathfrak{G}(u_1, u_2), \quad u_1, u_2 \in \mathbb{U},$$

which gives $R_\phi(t) \circ R_\phi^T(t) = \text{id}_\mathbb{V}$ and $R_\phi^T(t) \circ R_\phi(t) = \text{id}_\mathbb{U}$. Differentiating these relations with respect to t gives this part of the result.

(ii) Now suppose that \tilde{x}_0 and $\tilde{X}_0 = X_0$ are alternate origins for \mathbb{S} and \mathbb{B} . As we saw in the proof of Lemma 2.8, we have

$$\tilde{r}_\phi(t) = \phi(t)(X_0) - \tilde{x}_0 = \phi(t)(X_0) - x_0 - (\tilde{x}_0 - x_0) = r_\phi(t) - (\tilde{x}_0 - x_0).$$

Thus $\dot{\tilde{r}}_\phi(t) = \dot{r}_\phi(t)$. \blacksquare

The lemma has the following corollary that will be useful to us when representing velocities associated to rigid motions of a rigid body.

³In fact, two representations of $\text{Rgd}(\mathcal{B}; \mathcal{S})$ by $\text{SO}(3) \times \mathbb{R}^3$ are related by a *real analytic* diffeomorphism. Thus we can actually talk about regularity up to real analyticity. Indeed, it is perhaps most natural to speak of “real analyticity” rather than “smoothness” in mechanics, since there are probably no physical phenomenon that can be modelled smoothly that cannot also be modelled real analytically. However, burdening the presentation with real analyticity is not something we will undertake.

2.14 Corollary: (A convenient representation of the velocity of a rigid motion)

Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, and let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. Let $X_0 \in \mathbf{B}$ and $x_0 \in \mathbf{S}$. Then there is an injective vector bundle mapping

$$\tau_{X_0}: \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S}) \times (\text{Hom}_{\mathbb{R}}(\mathbf{U}; \mathbf{V}) \oplus \mathbf{V}),$$

depending only on X_0 and satisfying $\tau_{X_0}(\dot{\phi}(t)) = (\phi(t), (\dot{R}_\phi(t), \dot{r}_\phi(t)))$ for every differentiable rigid motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$.

Proof: That the mapping τ_{X_0} as defined depends only on X_0 follows since R_ϕ depends only on ϕ (by Lemma 2.8) and since \dot{r}_ϕ depends only on X_0 (by Lemma 2.13). To see that τ_{X_0} is injective, let

$$\begin{aligned} \Gamma: \text{Rgd}(\mathcal{B}; \mathcal{S}) \times \mathbf{B} &\rightarrow \mathbf{S} \\ (\Phi, X) &\mapsto \Phi(X) \end{aligned}$$

and, for $X \in \mathbf{B}$, define $\Gamma^X: \text{Rgd}(\mathcal{B}; \mathcal{S}) \rightarrow \mathbf{S}$ by $\Gamma^X(\Phi) = \Gamma(\Phi, X)$. In like manner define

$$\begin{aligned} \Gamma_{\text{body}}: \text{Rgd}(\mathcal{B}) \times \mathbf{B} &\rightarrow \mathbf{B} \\ (\Psi, X) &\mapsto \Psi(X) \end{aligned}$$

and $\Gamma_{\text{body}}^X(\Psi) = \Gamma_{\text{body}}(\Psi, X)$. Now suppose that $\tau_{X_0}(\dot{\phi}(t_0)) = 0$ for some $t_0 \in \mathbb{T}'$. Then $\dot{R}_\phi(t_0) = 0$ and $\dot{r}_\phi(t_0) = 0$. Since

$$\phi(t)(X) = x_0 + (r_\phi(t) + R_\phi(t)(X - X_0)),$$

it follows that $\frac{d}{dt}\big|_{t=t_0}(\phi(t)(X)) = 0$ for every $X \in \mathbf{B}$. Note that $\frac{d}{dt}(\phi(t)(X)) = T_{\phi(t)}\Gamma^X(\dot{\phi}(t))$ so that $\dot{\phi}(t_0) \in \ker(T_{\phi(t_0)}\Gamma^X)$ for every $X \in \mathbf{B}$. Now consider the identification $\Phi \mapsto \phi(t_0)^{-1} \circ \Phi$ of $\text{Rgd}(\mathcal{B}; \mathcal{S})$ with the group $\text{Rgd}(\mathcal{B})$, along with the following commutative diagram:

$$\begin{array}{ccc} \text{Rgd}(\mathcal{B}; \mathcal{S}) & \xrightarrow{\Gamma^X} & \mathbf{S} \\ \phi(t_0)^{-1} \circ \Phi \downarrow & & \downarrow \phi(t_0)^{-1} \\ \text{Rgd}(\mathcal{B}) & \xrightarrow{\Gamma_{\text{body}}^X} & \mathbf{B} \end{array}$$

This identifies \mathbf{S} with \mathbf{B} , and $\text{Rgd}(\mathcal{B}; \mathcal{S})$ with $\text{Rgd}(\mathcal{B})$ in such a way that the group action Γ is identified with the group action Γ_{body} . This group action is effective, and so $\bigcap_{X \in \mathbf{B}} T_{\text{id}}\Gamma_{\text{body}}^X = \{0\}$. This gives $\dot{\phi}(t_0) = 0$, giving injectivity of τ_{X_0} . \blacksquare

As an application of the preceding result, let us indicate the nature of the lift to $\mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$ of the natural left- and right-actions of $\text{Rgd}(\mathcal{S})$ and $\text{Rgd}(\mathcal{B})$.

2.15 Lemma: (The lifts of the natural group actions of the space of rigid transformations) Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model and let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space. Let $x_0 \in \mathbf{S}$ and $X_0 \in \mathbf{B}$. Consider the actions $\mathcal{A}_{\text{spatial}}$ and $\mathcal{A}_{\text{body}}$ of $\text{Rgd}(\mathcal{S})$ and $\text{Rgd}(\mathcal{B})$, respectively, on $\text{Rgd}(\mathcal{B}; \mathcal{S})$ (see Proposition 2.6). Then we have

$$T_{\Phi}\mathcal{A}_{\text{spatial}, \Psi}(A, v) = (\Psi \circ \Phi, (R_\Psi \circ A, R_\Psi(v))), \quad (A, v) \in \mathbb{T}_\Phi \text{Rgd}(\mathcal{B}; \mathcal{S}), \quad \Psi \in \text{Rgd}(\mathcal{S}),$$

and

$$T_{\Phi}\mathcal{A}_{\text{body}, \Psi}(A, v) = (\Phi \circ \Psi, (A \circ R_\Psi, v + A(r_\Psi))), \quad (A, v) \in \mathbb{T}_\Phi \text{Rgd}(\mathcal{B}; \mathcal{S}), \quad \Psi \in \text{Rgd}(\mathcal{B}).$$

Proof: We first prove the spatial formula. Given the choices of origin, we represent $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ with $(R_\Phi, r_\Phi) \in \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V}$ and $\Psi \in \text{Rgd}(\mathcal{S})$ with $(R_\Psi, r_\Psi) \in \text{SO}(\mathbf{V}, \mathfrak{g}, \theta) \times \mathbf{V}$. We then have that $\Psi \circ \Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ is represented by $(R_\Psi \circ R_\Phi, r_\Psi + R_\Psi(r_\Phi))$, cf. equation (2.3). Differentiating with respect to (R_Φ, r_Φ) gives the desired formula.

Similarly, for the body formula, we represent $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ with $(R_\Phi, r_\Phi) \in \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V}$ and $\Psi \in \text{Rgd}(\mathcal{B})$ with $(R_\Psi, r_\Psi) \in \text{SO}(\mathbf{U}, \mathfrak{G}, \Theta) \times \mathbf{U}$. We then have that $\Phi \circ \Psi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ is represented by $(R_\Phi \circ R_\Psi, r_\Phi + R_\Phi(r_\Psi))$. Differentiating with respect to (R_Φ, r_Φ) again gives the desired result. \blacksquare

We can now sensibly make the following definition.

2.16 Definition: (Rigid body velocities) Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathfrak{G}, \Theta)$ be a body reference space, let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval, and let $x_0 \in \mathbf{S}$ and $X_0 \in \mathbf{B}$. For a rigid motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ define $R_\phi: \mathbb{T}' \rightarrow \text{Isom}^+(\mathcal{B}; \mathcal{S})$ and $r_\phi: \mathbb{T}' \rightarrow \mathbf{V}$ by

$$\phi(t)(X) = x_0 + (r_\phi(t) + R_\phi(t)(X - X_0)), \quad t \in \mathbb{T}'.$$

We then make the following definitions.

(i) The *spatial angular velocity* for the motion is

$$t \mapsto \omega_\phi(t) \triangleq \widetilde{\dot{R}_\phi(t) R_\phi^T(t)} \in \mathbf{V}.$$

(ii) The *body angular velocity* for the motion is

$$t \mapsto \Omega_\phi(t) \triangleq \widetilde{R_\phi^T(t) \dot{R}_\phi(t)} \in \mathbf{U}.$$

(iii) The *spatial translational velocity* for the motion is

$$t \mapsto v_\phi(t) \triangleq \dot{r}_\phi(t) + r_\phi(t) \times \omega_\phi(t) \in \mathbf{V}.$$

(iv) The *body translational velocity* for the motion is

$$t \mapsto V_\phi(t) \triangleq R_\phi^T(\dot{r}_\phi(t)) \in \mathbf{U}. \quad \bullet$$

It is convenient to package these representations of velocity into vector bundle mappings

$$\tau_{\text{spatial}}: \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S}) \times (\mathbf{V} \oplus \mathbf{V})$$

and

$$\tau_{\text{body}}: \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S}) \times (\mathbf{U} \oplus \mathbf{U})$$

defined by

$$\tau_{\text{spatial}}(\dot{\phi}(t)) = (\phi(t), (\omega_\phi(t), v_\phi(t)))$$

and

$$\tau_{\text{body}}(\dot{\phi}(t)) = (\phi(t), (\Omega_\phi(t), V_\phi(t))),$$

for a differentiable mapping $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$. We note that the definition of τ_{spatial} requires a choice of both $x_0 \in \mathbb{S}$ and $X_0 \in \mathbb{B}$, whereas the definition of τ_{body} requires only a choice of $X_0 \in \mathbb{B}$.

With the possible exception of the spatial translational velocity (for a discussion of which we refer to [Murray, Li, and Sastry 1994, page 55]), these velocity representations have simple physical characterisations. The body translational velocity is the velocity of the image of the body reference point X_0 relative to the spatial reference point x_0 , viewed from the body coordinate system. The direction of the spatial angular velocity $\omega(t)$ is the axis about which the body instantaneously rotates at time t , seen in the spatial frame, while the body angular velocity is the same thing, seen in the body frame as it is mapped by the motion into the spatial frame. A thoughtful discussion of angular velocity is undertaken by Crampin [1986].

To close this section, we prove a couple of useful identities that relate to angular velocity.

2.17 Lemma: (Angular velocity identities) *Let $(\mathbb{U}, \mathbb{G}, \Theta)$ and $(\mathbb{V}, \mathbb{g}, \theta)$ be oriented three-dimensional inner product spaces and let $R \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$. Then*

- (i) $R(u_1 \times u_2) = (R(u_1)) \times (R(u_2))$ for $u_1, u_2 \in \mathbb{U}$ and
- (ii) $R \circ \hat{u} \circ R^T = \widehat{R(u)}$ for $u \in \mathbb{U}$.

Proof: (i) Because R is a linear isometry,

$$\mathbb{g}(R(u_1), R(u_2)) = \mathbb{G}(u_1, u_2), \quad \sqrt{\mathbb{g}(R(u_a), R(u_a))} = \sqrt{\mathbb{G}(u_a, u_a)}, \quad a \in \{1, 2\}.$$

Thus $\theta_{R(u_1), R(u_2)} = \theta_{u_1, u_2}$. Since R is a linear isometry, $R(n_{u,v})$ is orthogonal to $\text{span}(R(u_1), R(u_2))$. Since R is also orientation-preserving, $(R(u_1), R(u_2), R(n_{u_1, u_2}))$ is positively-oriented if (u_1, u_2, n_{u_1, u_2}) is. This shows that $R(n_{u_1, u_2}) = n_{R(u_1), R(u_2)}$ if u_1 and u_2 are not collinear. Then we have

$$\begin{aligned} (R(u_1)) \times (R(u_2)) &= \|R(u_1)\| \|R(u_2)\| \sin(\theta_{R(u_1), R(u_2)}) n_{R(u_1), R(u_2)} \\ &= \|u_1\| \|u_2\| \sin(\theta_{u_1, u_2}) R(n_{u_1, u_2}) = R(u_1 \times u_2), \end{aligned}$$

as desired.

(ii) For $u \in \mathbb{U}$ and $v \in \mathbb{V}$,

$$\hat{u} \circ R^T(v) = u \times (R^T(v)) = R^T(R(u) \times v),$$

using part (i), which gives

$$R \circ \hat{u} \circ R^T(v) = R(u) \times v.$$

From the definition of $\hat{\cdot}$, the result follows. ■

3. Rigid bodies

In this section we present a rigorous treatment of rigid bodies, and define the three attributes of a rigid body that are required to describe its dynamics.

3.1. Rigid bodies, mass, and centre of mass. Suppose that we have a body reference space $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$. Note that a body frame give us a representation of $X \in \mathbf{B}$ by

$$X = O_{\text{body}} + X^1 b_1 + X^2 b_2 + X^3 b_3$$

for unique $X^1, X^2, X^3 \in \mathbb{R}$. Thus we have, as with spatial frame, an identification of \mathbf{B} with \mathbb{R}^3 , and so we have a topology for \mathbf{B} induced by that for \mathbb{R}^3 . Moreover, rotation and translation invariance of the topology for \mathbb{R}^3 ensures that the topology for \mathbf{B} is independent of the choice of body frame. Moreover, this is also the metric topology for \mathbf{S} associated with the metric defined by \mathbf{G} , cf. (1.2). The point is that the ‘‘Borel σ -algebra’’ associated with \mathbf{B} is meaningful, and we denote it by $\mathcal{B}(\mathbf{B})$.

3.1 Definition: (Rigid body) Let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space. A *rigid body* in \mathcal{B} is a pair (\mathcal{B}, μ) , where $\mathcal{B} \subseteq \mathbf{B}$ is a compact set and μ is a finite positive Borel measure on \mathbf{B} with support equal to \mathcal{B} . •

The generality of the preceding definition allows us to consider particles as being special cases of rigid bodies. Indeed, a particle of mass m at $X_0 \in \mathbf{B}$ is, as a rigid body, $(\{X_0\}, m\delta_{X_0})$, where δ_{X_0} is the Dirac measure at X_0 defined by

$$\delta_{X_0}(B) = \begin{cases} 1, & X_0 \in B, \\ 0, & X_0 \notin B, \end{cases}$$

for any $B \in \mathcal{B}(\mathbf{B})$. The generality of the definition also allows for other idealisations of rigid bodies, such as bodies occupying a line segment in \mathbf{B} , i.e., infinitely thin rods. These sorts of constructions will lead us to a detailed consideration of degenerate rigid bodies in Section 4.

For a rigid body, we define the following two notions.

3.2 Definition: (Mass, centre of mass) Let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} .

- (i) The *mass* of (\mathcal{B}, μ) is $\mu(\mathcal{B})$.
- (ii) The *centre of mass* of (\mathcal{B}, μ) is

$$\mathbf{X}_c(\mathcal{B}, \mu) = X_0 + \frac{1}{\mu(\mathcal{B})} \int_{\mathcal{B}} (X - X_0) d\mu(X)$$

for $X_0 \in \mathbf{B}$. •

3.3 Remark: (Definition of centre of mass) The adding and subtracting of X_0 in the formula for the centre of mass is required since the integral is defined for *real-valued* functions, and so for finite-dimensional vector space-valued functions by choosing a basis. Thus the expression

$$\frac{1}{\mu(\mathcal{B})} \int_{\mathcal{B}} X d\mu(X)$$

is not defined, since the integrand is in an affine space. Thus this is made sense of by subtracting, then adding, X_0 . Moreover, the integral so defined is independent of the choice of X_0 , as we shall shortly see. •

When no ambiguity can arise, we shall denote the mass and centre of mass by m and X_c , respectively.

The introduction of a rigid body in a body reference space gives us an immediate physically meaningful origin, namely the centre of mass. We shall often make this choice without mention, particularly by using Corollary 2.14 to identify $\mathsf{T}(\mathsf{Rgd}(\mathcal{B}; \mathcal{S}))$ with a subset of $\mathsf{Rgd}(\mathcal{B}; \mathcal{S}) \times (\mathsf{Hom}_{\mathbb{R}}(\mathsf{U}; \mathsf{V}) \oplus \mathsf{V})$.

The following lemma records some useful properties of the centre of mass.

3.4 Lemma: (Properties of centre of mass) *Let $\mathcal{B} = (\mathsf{B}, \mathsf{U}, \mathsf{G}, \Theta)$ be a body reference space with (\mathcal{B}, μ) a rigid body. The following statements hold:*

(i) *the expression*

$$X_0 + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X - X_0) \, \mathrm{d}\mu(X)$$

is independent of X_0 ;

(ii) *X_c is the unique point in B with the property that*

$$\int_{\mathsf{B}} (X - X_c) \, \mathrm{d}\mu(X) = 0;$$

(iii) *$X_c \in \mathsf{int}_{\mathsf{aff}(\mathcal{B})}(\mathsf{conv}(\mathcal{B}))$.*

Proof: (i) Let $\tilde{X}_0 \in \mathsf{B}$ and compute

$$\begin{aligned} \tilde{X}_0 + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X - \tilde{X}_0) \, \mathrm{d}\mu(X) &= X_0 + (\tilde{X}_0 - X_0) + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X - X_0) \, \mathrm{d}\mu(X) \\ &\quad + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X_0 - \tilde{X}_0) \, \mathrm{d}\mu(X) \\ &= X_0 + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X - X_0) \, \mathrm{d}\mu(X). \end{aligned}$$

(ii) By definition of X_c and part (i), we have

$$X_c = X_c + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X - X_c) \, \mathrm{d}\mu(X),$$

from which it follows that

$$\int_{\mathsf{B}} (X - X_c) \, \mathrm{d}\mu(X) = \mu(\mathcal{B})(X_c - X_c) = 0.$$

Now suppose that $\tilde{X}_c \in \mathsf{B}$ has the property that

$$\int_{\mathsf{B}} (X - \tilde{X}_c) \, \mathrm{d}\mu(X) = 0.$$

Then, by (i),

$$X_c = \tilde{X}_c + \frac{1}{\mu(\mathcal{B})} \int_{\mathsf{B}} (X - \tilde{X}_c) \, \mathrm{d}\mu(X),$$

from which we conclude that $\tilde{X}_c = X_c$.

(iii) If X_c is on the relative boundary of $\text{conv}(\mathcal{B})$ or not in \mathcal{B} , then there exists a hyperplane P in \mathcal{B} passing through X_c such that there are points in \mathcal{B} which lie on one side of P , but there are no points in \mathcal{B} on the opposite side. In other words, there exists $\lambda \in \mathcal{U}^*$ such that the set

$$\{X \in \mathcal{B} \mid \lambda(X - X_c) \geq 0\}$$

is non-empty, but the set

$$\{X \in \mathcal{B} \mid \lambda(X - X_c) < 0\}$$

is empty. But this would imply that

$$\int_{\mathcal{B}} \lambda(X - X_c) d\mu(X) > 0,$$

contradicting (ii). ■

3.2. Inertia tensor. In the preceding definition we introduced two of the three components of a rigid body required to determine its dynamical behaviour. The third is the following.

3.5 Definition: (Inertia tensor) Let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $X_0 \in \mathbb{R}^3$. The *inertia tensor* of (\mathcal{B}, μ) about X_0 is the linear map $\mathbb{I}_{X_0}(\mathcal{B}, \mu) \in \text{Hom}_{\mathbb{R}}(\mathcal{U}; \mathcal{U})$ defined by

$$\mathbb{I}_{X_0}(\mathcal{B}, \mu)(u) = \int_{\mathcal{B}} (X - X_0) \times (u \times (X - X_0)) d\mu(X),$$

for $u \in \mathcal{U}$. By $\mathbb{I}_c(\mathcal{B}, \mu)$ we denote the inertia tensor about the centre of mass of (\mathcal{B}, μ) . •

When no ambiguity can arise, we shall denote the inertia tensor about X_0 by \mathbb{I}_{X_0} and the inertia tensor about the centre of mass by \mathbb{I}_c .

An essential physical property of the inertia tensor is the following.

3.6 Lemma: (Symmetry and definiteness of the inertia tensor) *For a body reference space $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$, for a rigid body (\mathcal{B}, μ) in \mathcal{B} , and for $X_0 \in \mathcal{B}$, \mathbb{I}_{X_0} is symmetric and positive-semidefinite with respect to the inner product \mathcal{G} .*

Proof: We prove the symmetry of \mathbb{I}_{X_0} as follows:

$$\begin{aligned} \mathcal{G}(\mathbb{I}_{X_0}(v_1), v_2) &= \int_{\mathcal{B}} \mathcal{G}((X - X_0) \times (v_1 \times (X - X_0)), v_2) d\mu(X) \\ &= \int_{\mathcal{B}} \mathcal{G}(v_1 \times (X - X_0), v_2 \times (X - X_0)) d\mu(X) \\ &= \int_{\mathcal{B}} \mathcal{G}(v_1, (X - X_0) \times (v_2 \times (X - X_0))) d\mu(X) \\ &= \mathcal{G}(v_1, \mathbb{I}_{X_0}(v_2)), \end{aligned} \tag{3.1}$$

using (1.1). That \mathbb{I}_{X_0} is positive-semidefinite follows directly from (3.1), taking $v_1 = v_2 = v$. ■

From the lemma, we immediately deduce that \mathbb{I}_{X_0} has real, nonnegative eigenvalues, which we call *principal inertias*, and a basis of eigenvectors, which we call *principal axes*. In Section 4 we shall carefully study the situation when some or all of the principal inertias are zero.

3.3. Fibre semimetrics associated with a rigid body. The physical data of a rigid body can be conglomerated into one geometric object, which will also have spatial and body representations. In this section we introduce these geometric constructions.

Let us begin with the definition, noting that a Riemannian semimetric has all of the attributes of a Riemannian metric, except it is only positive-semidefinite rather than positive-definite.

3.7 Definition: (Kinetic energy semimetric) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . The *kinetic energy semimetric* on $\text{Rgd}(\mathcal{B}; \mathcal{S})$ is the Riemannian semimetric $\mathbb{G}_{\mathbb{B}}$ defined by

$$\mathbb{G}_{\mathbb{B}}((A_1, v_1), (A_2, v_2)) = \mathbb{G}(\mathbb{I}_c(\widetilde{R_{\Phi}^T \circ A_1}), \widetilde{R_{\Phi}^T \circ A_2}) + mg(v_1, v_2)$$

for $(A_1, v_1), (A_2, v_2) \in \mathbb{T}_{\Phi}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$, cf. Corollary 2.14. •

The motivation for calling this the kinetic energy semimetric is given by Lemma 7.8 below.

Let us give a basic property of the kinetic energy semimetric.

3.8 Lemma: (Invariance of the kinetic energy semimetric under spatial symmetry) *Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . The kinetic energy semimetric $\mathbb{G}_{\mathbb{B}}$ is invariant under the left-action of $\text{Rgd}(\mathcal{S})$ on $\text{Rgd}(\mathcal{B}; \mathcal{S})$. (See Proposition 2.6 for the definition of the group action.)*

Proof: Recall that the left action in question is denoted by $\mathcal{A}_{\text{spatial}}$ and is defined by $\mathcal{A}_{\text{spatial}}(\Psi, \Phi) = \Psi \circ \Phi$. Let $\Psi \in \text{Rgd}(\mathcal{S})$ and note that

$$T_{\Phi} \mathcal{A}_{\text{spatial}, \Psi}(A, v) = (R_{\Psi} \circ A, R_{\Psi}(v))$$

by Lemma 2.15. We then directly compute

$$\mathbb{G}(\mathbb{I}_c((R_{\Psi} \circ R_{\Phi})^T \circ (R_{\Psi} \circ A_1), (R_{\Psi} \circ R_{\Phi})^T \circ (R_{\Psi} \circ A_2))) = \mathbb{G}(\mathbb{I}_c(\widetilde{R_{\Phi}^T \circ A_1}), \widetilde{R_{\Phi}^T \circ A_2})$$

and

$$mg(R_{\Psi}(v_1), R_{\Psi}(v_2)) = mg(v_1, v_2),$$

the latter since R_{Ψ} is \mathfrak{g} -orthogonal. These two formulae together show that $\mathbb{G}_{\mathbb{B}}$ is $\text{Rgd}(\mathcal{S})$ -invariant. ■

Let us give the representation of $\mathbb{G}_{\mathbb{B}}$ in spatial and body velocity representations. Thus we define fibre semimetrics $\mathbb{G}_{\mathbb{B}}^{\text{spatial}}$ and $\mathbb{G}_{\mathbb{B}}^{\text{body}}$ on the vector bundles $\text{Rgd}(\mathcal{B}; \mathcal{S}) \times (\mathbb{V} \oplus \mathbb{V})$ and $\text{Rgd}(\mathcal{B}; \mathcal{S}) \times (\mathbb{U} \oplus \mathbb{U})$, respectively, by

$$\mathbb{G}_{\mathbb{B}}^{\text{spatial}}((\omega_1, v_1), (\omega_2, v_2)) = \mathbb{G}_{\mathbb{B}}(\tau_{\text{spatial}}^{-1}(\omega_1, v_1), \tau_{\text{spatial}}^{-1}(\omega_2, v_2)) \quad (3.2)$$

and

$$\mathbb{G}_{\mathbb{B}}^{\text{body}}((\Omega_1, V_1), (\Omega_2, V_2)) = \mathbb{G}_{\mathbb{B}}(\tau_{\text{body}}^{-1}(\Omega_1, V_1), \tau_{\text{body}}^{-1}(\Omega_2, V_2)), \quad (3.3)$$

respectively. Let us give the explicit expression for these semimetrics.

3.9 Lemma: (Spatial and body representations of the kinetic energy semimetric)

Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $x_0 \in \mathbf{S}$. By Lemma 2.8, associate $(R_\Phi, r_\Phi) \in \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V}$ to $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$. We then have

$$\mathbf{G}_{\mathcal{B}}^{\text{spatial}}((\omega_1, v_1), (\omega_2, v_2)) = \mathbf{G}(\mathbb{I}_c(R_\Phi^T(\omega_1)), R_\Phi^T(\omega_2)) + m\mathbf{g}(v_1 - r_\Phi \times \omega_1, v_2 - r_\Phi \times \omega_2)$$

for $(\omega_1, v_1), (\omega_2, v_2)$ in the fibre of $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, and

$$\mathbf{G}_{\mathcal{B}}^{\text{body}}((\Omega_1, V_1), (\Omega_2, V_2)) = \mathbf{G}(\mathbb{I}_c(\Omega_1), \Omega_2) + m\mathbf{G}(V_1, V_2)$$

for $(\Omega_1, V_1), (\Omega_2, V_2)$ in the fibre of $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$.

Proof: Note that $\hat{\omega} = R_\phi \circ \hat{\Omega} \circ R_\Phi^T$. The result follows from applying this fact, along with Lemma 2.17, for the first equation and the fact that $R_\Phi \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$ for the second equation. \blacksquare

4. Rigid bodies with degenerate inertia tensors

One of the contributions of our approach to rigid body dynamics is that we carefully consider the situation of rigid bodies with degenerate inertia tensor, i.e., particles and infinitely thin rods. We shall see that, in such cases, the inertia tensor possesses an internal symmetry that we will use to perform a reduction of the space $\text{Rgd}(\mathcal{B}; \mathcal{S})$ of configurations of the rigid body. This, in turn, will be essential in our construction of the usual “kinetic energy” of a mechanical system that is essential in the formulation of the Euler–Lagrange equations.

4.1. Degeneracies of the inertia tensor and internal symmetry. Let us first consider the possible ways in which the inertia tensor can be degenerate.

4.1 Lemma: (Degenerate inertia tensors) Let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $X_0 \in \mathbf{B}$. Let \mathbb{I}_{X_0} denote the inertia tensor of (\mathcal{B}, μ) about X_0 . The following statements hold:

- (i) if \mathbb{I}_{X_0} has a zero eigenvalue, then the other two eigenvalues are equal;
- (ii) if \mathbb{I}_{X_0} has two zero eigenvalues, then $\mathbb{I}_{X_0} = 0$.

Proof: (i) Let u be a unit eigenvector for the zero eigenvalue. We claim that the support of the measure μ must be contained in the line

$$\ell_u = \{X_0 + su \mid s \in \mathbb{R}\}.$$

To see that this must be so, suppose that the support of μ is not contained in ℓ_u . Then there exists a Borel set $B \subseteq \mathbf{B} \setminus \ell_u$ so that $\mu(B) > 0$. This would imply that

$$\begin{aligned} \mathbb{I}_{X_0}(u, u) &= \int_{\mathcal{B}} \mathbf{G}(u \times (X - X_0), u \times (X - X_0)) \, d\mu(X) \\ &\geq \int_B \mathbf{G}(u \times (X - X_0), u \times (X - X_0)) \, d\mu(X). \end{aligned}$$

Since $B \cap \ell_u = \emptyset$, it follows that, for all points $X \in B$, the vector $X - X_0$ is not collinear with u . Therefore

$$\mathbb{G}(u \times (X - X_0), u \times (X - X_0)) > 0$$

for all $X \in B$, and this would imply that $\mathbb{I}_{X_0}(u, u) > 0$. But this contradicts u being an eigenvector with zero eigenvalue, and so the support of μ must be contained in the line ℓ_u .

To see that this implies that the other eigenvalues are equal, we shall show that any vector that is \mathbb{G} -orthogonal to u is an eigenvector for \mathbb{I}_{X_0} . First let (u_1, u_2) be an orthonormal basis for the \mathbb{G} -orthogonal complement to $\text{span}(u)$ and write

$$X - X_0 = f(X)u + g^1(X)u_1 + g^2(X)u_2$$

for functions $f, g^1, g^2 : \mathbb{B} \rightarrow \mathbb{R}$. Since the support of μ is contained in the line ℓ_u , we have

$$\int_{\mathbb{B}} (X - X_0) \times (v \times (X - X_0)) d\mu(X) = u \times (v \times u) \int_{\mathbb{B}} (f(X))^2 d\mu(X)$$

for all $v \in \mathbb{U}$. Now recall the property of the cross-product that $u \times (v \times u) = v$ provided v is orthogonal to u and that u has unit length. Therefore, we see that, for any v that is orthogonal to u , we have

$$\mathbb{I}_{X_0}(v) = \left(\int_{\mathbb{B}} (f(X))^2 d\mu(X) \right) v,$$

meaning that all such vectors v are eigenvectors with the same eigenvalue, which is what we wished to show.

(ii) It follows from our above arguments that, if the zero eigenvalue has multiplicity 2, then the support of μ must lie in the intersection of lines ℓ_{u_1} and ℓ_{u_2} for orthogonal eigenvectors for the zero eigenvalue. This intersection is a single point that must, therefore, be X_0 . From this and the definition of \mathbb{I}_{X_0} it follows that $\mathbb{I}_{X_0} = 0$. \blacksquare

Note that in, proving the result, we have proven the following corollary.

4.2 Corollary: (The ‘‘shape’’ of rigid bodies with degenerate inertia tensors) Let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $X_0 \in \mathbb{B}$. Let \mathbb{I}_{X_0} denote the inertia tensor of (\mathcal{B}, μ) about X_0 . The following statements hold:

- (i) \mathbb{I}_{X_0} has a zero eigenvalue if and only if \mathcal{B} is contained in a line through X_0 ;
- (ii) if \mathbb{I}_{X_0} has two zero eigenvalues then $\mathcal{B} = \{X_0\}$, i.e., \mathcal{B} is a particle located at x_0 ;
- (iii) if there is no line through X_0 that contains the support of μ , then the inertia tensor is nondegenerate.

With the preceding discussion undertaken, we make the following essential definition.

4.3 Definition: (Internal symmetry group of a rigid body) Let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . The *internal symmetry group* of (\mathcal{B}, μ) is the subgroup $\mathbb{I}(\mathcal{B})$ of $\text{Rgd}(\mathcal{B})$ with the property that $\Phi \in \mathbb{I}(\mathcal{B})$ if and only if $\Phi(\mathcal{B}') = \mathcal{B}'$ for any rigid body (\mathcal{B}', μ') in \mathcal{B} for which $\mu'(\mathcal{B}') = \mu(\mathcal{B})$, $X_c(\mathcal{B}', \mu') = X_c(\mathcal{B}, \mu)$, and $\mathbb{I}_c(\mathcal{B}', \mu') = \mathbb{I}_c(\mathcal{B}, \mu)$. \bullet

Let us illustrate the intent of this definition by considering some examples.

4.4 Examples: (Internal symmetry groups)

1. Let us consider a particle of mass m located at $X_0 \in \mathcal{B}$ as a rigid body, i.e., $(\mathcal{B}, \mu) = (\{X_0\}, m\delta_{X_0})$. Here we claim that $\Phi \in \mathcal{I}(\mathcal{B})$ if and only if $\Phi(X) = R(X - X_0)$ for some $R \in \text{SO}(\mathcal{U}, \mathcal{G}, \Theta)$, i.e., $\mathcal{I}(\mathcal{B})$ consists of rotations about X_0 .

First we note that, if (\mathcal{B}', μ') is such that $\mu'(\mathcal{B}') = \mu(\mathcal{B})$, $X_c(\mathcal{B}', \mu') = X_c(\mathcal{B}, \mu)$, and $\mathbb{I}_c(\mathcal{B}', \mu') = \mathbb{I}_c(\mathcal{B}, \mu)$, then $(\mathcal{B}', \mu') = (\mathcal{B}, \mu)$ because of the fact that $\text{supp}(\mu) = \{X_0\}$. Clearly then, if Φ is a rotation about X_0 , then $\Phi \in \mathcal{I}(\mathcal{B})$.

Conversely, if $\Phi \in \mathcal{I}(\mathcal{B})$, then $\Phi(X_0) = X_0$ since $\mathcal{B} = \{X_0\}$. Thus Φ must be a rotation about X_0 .

2. Next let us consider the case where (\mathcal{B}, μ) is a rigid body, and not a particle, for which the support of μ is contained in a line

$$\ell_u = \{X_0 + su \mid s \in \mathbb{R}\}$$

for some nonzero $u \in \mathcal{U}$. By Lemma 3.4(iii) we may take $X_0 = X_c(\mathcal{B}, \mu)$. We claim that

$$\mathcal{I}(\mathcal{B}) = \{\Phi \in \text{Rgd}(\mathcal{B}) \mid \Phi(X) = R(X - X_c(\mathcal{B}, \mu)), R \in \text{SO}(\mathcal{U}, \mathcal{G}, \Theta), R(u) = u\},$$

i.e., $\mathcal{I}(\mathcal{B})$ consists of rotations about the u -axis fixing $X_c(\mathcal{B}, \mu)$.

First suppose that Φ is such a mapping. Let (\mathcal{B}', μ') have the same mass, centre of mass, and inertia tensor as (\mathcal{B}, μ) . Then the support of μ' is contained in ℓ_u by Corollary 4.2. Since $\Phi(X) = X$ for any $X \in \ell_u$, it follows that $\Phi(\mathcal{B}') = \mathcal{B}'$.

If $\Phi \in \mathcal{I}(\mathcal{B})$ then $\Phi(X_c(\mathcal{B}, \mu)) = X_c(\mathcal{B}, \mu)$ because $\Phi(\mathcal{B}) = \mathcal{B}$. Thus $\Phi(X) = R(X - X_c(\mathcal{B}, \mu))$ for some $R \in \text{SO}(\mathcal{U}, \mathcal{G}, \Theta)$. Since $\mathcal{B} \subseteq \ell_u$ is not a particle, there exists $X \in \mathcal{B} \setminus X_c(\mathcal{B}, \mu)$. Then $R(X - X_c(\mathcal{B}, \mu)) = X - X_c(\mathcal{B}, \mu)$. Since $X - X_c(\mathcal{B}, \mu) = au$ for some $a \neq 0$ we have

$$R(au) = au \implies R(u) = u,$$

i.e., Φ is a rotation about u fixing X_0 .

3. Next we consider a rigid body (\mathcal{B}, μ) with nondegenerate inertia tensor. Here we claim that $\mathcal{I}(\mathcal{B}) = \{\text{id}_{\mathcal{B}}\}$. Since it is clear that $\{\text{id}_{\mathcal{B}}\} \subseteq \mathcal{I}(\mathcal{B})$, it is only the other inclusion that requires proof.

Let $I_1, I_2, I_3 \in \mathbb{R}_{>0}$ be the principal inertias for (\mathcal{B}, μ) and let $\{u_1, u_2, u_3\}$ be principal axes. We abbreviate $m = \mu(\mathcal{B})$ and $X_c = X_c(\mathcal{B}, \mu)$. Denote

$$\begin{aligned} \text{R-M} = \{ & ((r_1, r_2, r_3), (m_1, m_2, m_3)) \in \mathbb{R}_{>0}^3 \times \mathbb{R}_{>0}^3 \mid \\ & 2(m_1 + m_2 + m_3) = m, m_j r_j^2 = I_j, j \in \{1, 2, 3\}\}. \end{aligned}$$

Choose

$$((r_1, r_2, r_3), (m_1, m_2, m_3)) \in \text{R-M}$$

so that r_1, r_2 , and r_3 are distinct. Denote

$$\mathcal{B}' = \{X_c + r_1 u_1\} \cup \{X_c - r_1 u_1\} \cup \{X_c + r_2 u_2\} \cup \{X_c - r_2 u_2\} \cup \{X_c + r_3 u_3\} \cup \{X_c - r_3 u_3\}.$$

We define a measure μ' with support equal to \mathcal{B}' by asking that μ' be a sum of six Dirac measures:

$$\mu' = \sum_{j=1}^3 (m_j \delta_{X_c + r_j u_j} + m_j \delta_{X_c - r_j u_j}).$$

One readily verifies that (\mathcal{B}', μ') has the same mass, centre of mass, and inertia tensor as (\mathcal{B}, μ) (this is by definition of R-M). If $\Phi \in \mathfrak{l}(\mathcal{B})$, then $\Phi(X_c) = X_c$ (again, because $\Phi(\mathcal{B}) = \mathcal{B}$) and, since $\Phi(\mathcal{B}') = \mathcal{B}'$ and r_1, r_2 , and r_3 are distinct,

$$\Phi(X_c \pm r_j u_j) \in \{X_c + r_j u_j, X_c - r_j u_j\}, \quad j \in \{1, 2, 3\}.$$

That is, Φ is a rotation fixing X_c about the axes u_1, u_2 , and u_3 . The only such rotation is the identity. •

Note that, by Lemma 4.1, the preceding examples suffice to cover all possible rigid bodies. The examples also show that $\mathfrak{l}(\mathcal{B})$ is a Lie subgroup of $\text{Rgd}(\mathcal{B})$. Moreover, the examples also give the following simple result which we shall have occasion to use.

4.5 Lemma: (Points fixed by the internal symmetry group) *Let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$ be a body reference space and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Then*

$$\{X \in \mathcal{B} \mid \Psi(X) = X \text{ for all } \Psi \in \mathfrak{l}(\mathcal{B})\} = \text{aff}(\mathcal{B}).$$

Next we consider the invariance of the inertia tensor under the internal symmetry group. To do this, we note that elements of the internal symmetry group are rotations about X_c , and so we can regard $\mathfrak{l}(\mathcal{B})$ as a Lie subgroup of $\text{SO}(\mathcal{U}; \mathcal{G}, \Theta)$. For $\Phi \in \mathfrak{l}(\mathcal{B})$ we shall denote by $R_\Phi \in \text{SO}(\mathcal{U}, \mathcal{G}, \Theta)$ the corresponding rotation. With this notation, we have the following result.

4.6 Lemma: (Invariance of inertia tensor under internal symmetries) *Let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$ be a body reference space and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . For each $\Psi \in \mathfrak{l}(\mathcal{B})$, we have $R_\Psi \circ \mathbb{I}_c \circ R_\Psi^T = \mathbb{I}_c$.*

Proof: Note that

$$\Psi(X) = X_c + R_\Psi(X - X_c) \quad \implies \quad \Psi^{-1}(X) - X_c = R_\Psi^T(X - X_c).$$

We then compute, using Lemma 4.5,

$$\begin{aligned} \mathbb{I}_c \circ R_\Psi(u) &= \int_{\mathcal{B}} (X - X_c) \times (R_\Psi(u) \times (X - X_c)) \, d\mu(X) \\ &= R_\Psi \int_{\mathcal{B}} (R_\Psi^T(X - X_c) \times (u \times R_\Psi^T(X - X_c))) \, d\mu(X) \\ &= R_\Psi \int_{\mathcal{B}} (\Psi^{-1}(X) - X_c) \times (u \times (\Psi^{-1}(X) - X_c)) \, d\mu(X) \\ &= R_\Psi \int_{\mathcal{B}} (X - X_c) \times (u \times (X - X_c)) \, d\mu(X) \\ &= R_\Psi \circ \mathbb{I}_c(u), \end{aligned}$$

as desired. ■

We then have the following more or less immediate consequence for the kinetic energy semimetric.

4.7 Corollary: (Invariance of the kinetic energy semimetric under internal symmetry) *Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . The kinetic energy semimetric $\mathbb{G}_{\mathcal{B}}$ is invariant under the right-action of $\mathbb{I}(\mathcal{B})$ on $\text{Rgd}(\mathcal{B}; \mathcal{S})$. (See Proposition 2.6 for the definition of the group action.)*

Proof: By Example 4.4 we have that $\mathbb{I}(\mathcal{B})$ consists solely of orientation-preserving rotations of \mathbf{B} that fix X_c . For $\Psi \in \mathbb{I}(\mathcal{B})$, let $R_{\Psi} \in \text{SO}(\mathbf{U}, \mathbf{G}, \Theta)$ be as in Lemma 4.6. For $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ and for $A_1, A_2 \in \mathbb{T}_{R_{\Phi}} \text{Isom}^+(\mathcal{B}; \mathcal{S})$, we then compute

$$\begin{aligned} & \mathbb{G}(\mathbb{I}_c(\overline{(R_{\Phi} \circ R_{\Psi})^T \circ (A_1 \circ R_{\Psi})}), \overline{(R_{\Phi} \circ R_{\Psi})^T \circ (A_2 \circ R_{\Psi})}) \\ &= \mathbb{G}(\mathbb{I}_c(R_{\Psi}^T \circ \overline{R_{\Phi}^T \circ (A_1 \circ R_{\Psi})}), \overline{R_{\Psi}^T \circ R_{\Phi}^T \circ (A_2 \circ R_{\Psi})}) \\ &= \mathbb{G}(\mathbb{I}_c \circ R_{\Psi}^T(\overline{R_{\Phi}^T \circ A_1}), \overline{R_{\Psi}^T \circ R_{\Phi}^T \circ A_2}) \\ &= \mathbb{G}(R_{\Psi} \circ \mathbb{I}_c \circ \overline{R_{\Psi}^T \circ (R_{\Phi}^T \circ A_1)}, \overline{R_{\Phi}^T \circ A_2}) \\ &= \mathbb{G}(\mathbb{I}_c(\overline{R_{\Phi}^T \circ A_1}), \overline{R_{\Phi}^T \circ A_2}), \end{aligned}$$

using Lemma 2.17, the fact that $R_{\Psi} \in \text{SO}(\mathbf{U}, \mathbf{G}, \Theta)$, and Lemma 4.6. This is, by Lemma 2.15, exactly the assertion that $\mathbb{G}_{\mathcal{B}}$ is invariant under $\mathbb{I}(\mathcal{B})$. \blacksquare

The Lie subalgebra associated to the Lie subgroup $\mathbb{I}(\mathcal{B})$ we denote by $\mathfrak{i}(\mathcal{B}, \mu)$. Just as we identify $\mathbb{I}(\mathcal{B})$ with a Lie subgroup of $\text{SO}(\mathbf{U}, \mathbf{G}, \Theta)$, we identify $\mathfrak{i}(\mathcal{B}, \mu)$ with a Lie subalgebra of $\mathfrak{so}(\mathbf{U}, \mathbf{G}, \Theta)$. For $\xi \in \mathfrak{i}(\mathcal{B}, \mu)$, we denote by $\hat{\Omega}_{\xi} \in \mathfrak{so}(\mathbf{U}, \mathbf{G}, \Theta)$ the associated skew-symmetric linear map, with $\Omega_{\xi} \in \mathbf{U}$ the associated vector. The following result gives a characterisation of the Lie algebra $\mathfrak{i}(\mathcal{B}, \mu)$ in terms of the inertia tensor.

4.8 Lemma: (The Lie algebra of the internal symmetry group) *Let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Then we have*

$$\ker(\mathbb{I}_c) = \{\Omega_{\xi} \mid \xi \in \mathfrak{i}(\mathcal{B}, \mu)\}.$$

Proof: We prove this case-by-case, following Example 4.4. In the case when $\dim(\mathfrak{i}(\mathcal{B}, \mu)) = 0$, i.e., when the inertia tensor is nondegenerate, then we have

$$\ker(\mathbb{I}_c) = \{0\} = \{\Omega_{\xi} \mid \xi \in \mathfrak{i}(\mathcal{B}, \mu)\}.$$

In the case when $\dim(\mathfrak{i}(\mathcal{B}, \mu)) = 3$, then the body is a particle and

$$\ker(\mathbb{I}_c) = \mathbf{U} = \{\Omega_{\xi} \mid \xi \in \mathfrak{i}(\mathcal{B}, \mu)\},$$

the latter equality because a particle has the property that

$$\{R_{\Phi} \mid \Phi \in \mathbb{I}(\mathcal{B})\} = \text{SO}(\mathbf{U}, \mathbf{G}, \Theta).$$

Finally, we consider the case $\dim(\mathfrak{i}(\mathcal{B}, \mu)) = 1$. In this case, the measure μ has support contained in a line $\{X_c + su \mid s \in \mathbb{R}\}$ for some unit vector $u \in \mathbf{U}$, and $\mathbb{I}(\mathcal{B})$ consists of the rotations about this line. Thus

$$\{\hat{\Omega}_{\xi} \mid \xi \in \mathfrak{i}(\mathcal{B}, \mu)\} = \text{span}(\hat{u}).$$

However, as we saw in the proof of Lemma 4.1, we have $\ker(\mathbb{I}_c) = \text{span}(u)$, giving the result. \blacksquare

4.2. Reduced configuration space. In the development of our correspondence of the Newton–Euler equations with the Euler–Lagrange equations, we will be careful with degenerate inertia tensors. To do this, we will quotient by the action of the internal symmetry group, and in this section we consider this process for a single rigid body.

We consider a Newtonian space model $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathbf{g}, \theta)$ and a body reference space $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$, with (\mathbb{B}, μ) a rigid body in \mathcal{B} . As usual, $\mathbb{I}(\mathbb{B})$ denotes the internal symmetry group of the body, and this acts freely and properly on the right on $\text{Rgd}(\mathcal{B}; \mathcal{S})$ by

$$\text{Rgd}(\mathcal{B}; \mathcal{S}) \times \mathbb{I}(\mathbb{B}) \ni (\Phi, \Psi) \mapsto \Phi \circ \Psi \in \text{Rgd}(\mathcal{B}; \mathcal{S}).$$

If we denote by

$$\mathbb{S}(\mathbb{V}, \mathbf{g}) = \{v \in \mathbb{V} \mid \mathbf{g}(v, v) = 1\}$$

the unit sphere in \mathbb{V} , then we have the following result describing the orbit space of this group action.

4.9 Proposition: (Reduced configurations of a rigid body) *Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . There is a one-to-one correspondence between $\text{Rgd}(\mathcal{B}; \mathcal{S})/\mathbb{I}(\mathbb{B})$ and*

- (i) $\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V}$ (when \mathbb{I}_c is nondegenerate),
- (ii) $\mathbb{S}(\mathbb{V}, \mathbf{g}) \times \mathbb{V}$ (when \mathbb{I}_c is degenerate but nonzero), or
- (iii) \mathbb{V} (when \mathbb{I}_c is zero).

Moreover, the canonical projection $\pi_{\mathcal{B}}: \text{Rgd}(\mathcal{B}; \mathcal{S}) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})/\mathbb{I}(\mathbb{B})$ is a smooth surjective submersion.

Proof: As we have seen, if $\Psi \in \mathbb{I}(\mathbb{B})$ then $\Psi(X_c) = X_c$. Thus Ψ is an orientation-preserving rotation about X_c . Thus we can write

$$\Psi(X) = X_c + R_{\Psi}(X - X_c)$$

for some $R_{\Psi} \in \text{SO}(\mathbb{U}, \mathbb{G}, \Theta)$. By choosing $x_0 \in \mathbb{S}$, for $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ we can write

$$\Phi(X) = x_0 + (r_{\Phi} + R_{\Phi}(X - X_c))$$

for some $r_{\Phi} \in \mathbb{V}$ and $R_{\Phi} \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$. We now consider the three cases.

The first case we consider is when the inertia tensor is nondegenerate. In this case, $\mathbb{I}(\mathbb{B})$ is the identity subgroup, and the result follows by assigning to the orbit through Φ the element $(R_{\Phi}, r_{\Phi}) \in \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V}$.

Next we consider the case when \mathbb{I}_c is degenerate, but nonzero. Here, as we saw in Example 4.4–2, $\mathbb{I}(\mathbb{B})$ consists of rotations about some $u \in \mathbb{U} \setminus \{0\}$ fixing X_c . We suppose that $\mathbb{G}(u, u) = 1$. To $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ we assign the element $(v_{\Phi}, r_{\Phi}) \in \mathbb{S}(\mathbb{V}, \mathbf{g}) \times \mathbb{V}$ by asking that

$$\Phi(X_c + au) = x_0 + (r_{\Phi} + av_{\Phi}), \quad a \in \mathbb{R}.$$

To show that this mapping is well-defined, we note that $R_{\Phi}(u) \in \mathbb{S}(\mathbb{V}, \mathbf{g})$ since R_{Φ} is a linear isometry. We also claim that the assignment establishes a bijective correspondence between $\text{Rgd}(\mathcal{B}; \mathcal{S})/\mathbb{I}(\mathbb{B})$ and $\mathbb{S}(\mathbb{V}, \mathbf{g}) \times \mathbb{V}$. To show injectivity, suppose that $(r_{\Phi_1}, v_{\Phi_1}) = (r_{\Phi_2}, v_{\Phi_2})$. Then $\Phi_1(X_c + au) = \Phi_2(X_c + au)$ which means that $\Phi_2^{-1} \circ \Phi_1$ fixes $X_c + au$ for every $a \in \mathbb{R}$. This means that $\Phi_2^{-1} \circ \Phi_1 \in \mathbb{I}(\mathbb{B})$ and so Φ_1 and Φ_2 are in the same orbit of the

action of $l(\mathcal{B})$. To prove surjectivity of the correspondence, let $(v, r) \in \mathfrak{S}(\mathbb{V}, \mathfrak{g}) \times \mathbb{V}$ and let $R \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$ satisfy $R(u) = v$. Then, if we take $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ to be defined by

$$\Phi(X) = x_0 + (r + R(X - X_c)),$$

we see that the orbit of Φ is associated to (v, r) by our correspondence.

Finally, if \mathbb{I}_c is zero, then $l(\mathcal{B})$ consists of all orientation preserving rotations about X_c . In this case, to $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ we assign $r_\Phi \in \mathbb{V}$ by asking that $\Phi(X_c) = x_0 + r_\Phi$. First we show that the correspondence between $\text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B})$ and \mathbb{V} is injective. Suppose that $r_{\Phi_1} = r_{\Phi_2}$. Then $\Phi_1(X_c) = \Phi_2(X_c)$ and so $\Phi_2^{-1} \circ \Phi_1$ is a rotation about X_c , giving that Φ_1 and Φ_2 are in the same orbit. For surjectivity of the correspondence, for $r \in \mathbb{V}$ we note that, if $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ satisfies

$$\Phi(X) = x_0 + (r + R(X - X_c))$$

for some $R \in \text{Isom}^+(\mathcal{B}; \mathcal{S})$, then the orbit of Φ is associated with $r \in \mathbb{V}$.

The final assertion of the lemma follows from the following observation. By fixing $x_0 \in \mathbb{S}$ and $X_c \in \mathbb{S}$ and by fixing orthonormal spatial and body frames, we have a smooth diffeomorphism of $\text{Rgd}(\mathcal{B}; \mathcal{S})$ with $\text{SO}(3) \times \mathbb{R}^3$. The three cases of the internal symmetry groups then correspond to the subgroups $\{\text{id}_{\mathbb{R}^3}\} \times \{\mathbf{0}\}$, $\text{SO}(2) \times \{\mathbf{0}\}$, and $\text{SO}(3) \times \{\mathbf{0}\}$ of $\text{SO}(3) \times \mathbb{R}^3$, and their standard right-actions. The assertion then follows from [Abraham and Marsden 1978, Corollary 4.1.21]. \blacksquare

We remark that the correspondence of the lemma depends on the choice of $x_0 \in \mathbb{S}$, as well as the natural fixing of the point $X_c \in \mathbb{B}$.

Let us make a couple of observations about the various degenerate rigid bodies.

4.10 Remarks: (Reduced configuration spaces for degenerate rigid bodies)

1. Let $\text{Trans}(\mathbb{S})$ be the affine translation group of \mathbb{S} . Thus an element of $\text{Trans}(\mathbb{S})$ is a mapping of \mathbb{S} of the form $\tau_v: x \mapsto x + v$ for some $v \in \mathbb{V}$. Then we have a left-action of $\text{Trans}(\mathbb{S})$ on $\text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B})$ by

$$(\tau_v, [\Phi]) \mapsto [\tau_v \circ \Phi].$$

It is a simple matter to deduce that this action is well-defined and free. Moreover, in the case that $l(\mathcal{B}) \simeq \text{SO}(\mathbb{U}, \mathbb{G}, \Theta)$, we have that this action is transitive. This implies that the reduced configuration space in this case is identified with the set of spatial affine translations. This corresponds to what we expect for the configurations of a particle. Moreover, this space is then realised in two ways, one as a quotient by the internal symmetry group of the body and the other as an inclusion of the spatial affine translation group. That this quotient and inclusion agree is a consequence of the fact that the rotations about a point in the body reference space are a normal subgroup of $\text{Rgd}(\mathcal{B})$.

2. For rigid bodies with degenerate but nonzero inertia tensors, the situation is quite different from that for particles, in the sense that the reduced configuration space cannot be embedded as a submanifold of $\text{Rgd}(\mathcal{B}; \mathcal{S})$. To make this statement sensible (and correct), let us think for a moment about what reasonable properties such an embedding should have. Thus we assume that we have an embedding

$$\iota: \text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B}) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S}).$$

For such an embedding to be physically meaningful, it needs to interact with the fact that $\text{Rgd}(\mathcal{B}; \mathcal{S})$ corresponds to actual configurations of the body, as does the projection $\pi_{\mathcal{B}}$ onto the reduced configuration space. Therefore, for the embedding ι to respect the physically meaningful data, the following diagram should commute:

$$\begin{array}{ccccc} \text{Rgd}(\mathcal{B}; \mathcal{S})/I(\mathcal{B}) & \xrightarrow{\iota} & \text{Rgd}(\mathcal{B}; \mathcal{S}) & \xrightarrow{\pi_{\mathcal{B}}} & \text{Rgd}(\mathcal{B}; \mathcal{S})/I(\mathcal{B}) \\ & & \searrow & \swarrow & \\ & & \text{id} & & \end{array}$$

This, of course, means that ι is a section of $\pi_{\mathcal{B}}$. However, no continuous section of $\pi_{\mathcal{B}}$ exists. This we argue as follows. First, let us make everything concrete by choosing spatial and body frames so that what we have is a continuous section of

$$\pi_{\mathcal{B}}: \text{SO}(3) \times \mathbb{R}^3 \rightarrow \mathbb{S}^2 \times \mathbb{R}^3.$$

By restriction, this implies that we have a continuous section of $\text{SO}(3) \rightarrow \mathbb{S}^2$. By [Husemoller 1994, Corollary 4.8.3] this implies that the bundle $\text{SO}(3) \rightarrow \mathbb{S}^2$ is trivial. This would imply that $\text{SO}(3)$ is homeomorphic to $\mathbb{S}^2 \times \mathbb{S}^1$. This, however, is not the case, since the fundamental group of $\text{SO}(3)$ is \mathbb{Z}_2 [Hatcher 2002, Example 1.43 and Page 291] and the fundamental group of $\mathbb{S}^2 \times \mathbb{S}^1$ is \mathbb{Z} [Hatcher 2002, Theorem 1.7, and Propositions 1.12 and 1.14]. This shows that the reduction process we use in the paper for degenerate bodies cannot be replaced in a rational and general way by using submanifolds of $\text{Rgd}(\mathcal{B}; \mathcal{S})$. •

According to the previous discussion, we make the following definition.

4.11 Definition: (Reduced configuration space of a rigid body) Let $\mathcal{S} = (\mathcal{S}, \mathcal{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . The *reduced configuration space* for (\mathcal{B}, μ) is $\text{Rgd}(\mathcal{B}; \mathcal{S})/I(\mathcal{B})$. By $\pi_{\mathcal{B}}: \text{Rgd}(\mathcal{B}; \mathcal{S}) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})/I(\mathcal{B})$ we denote the canonical projection. •

4.3. Reduced velocities. The following result gives the form of the vertical bundle for the surjective submersion $\pi_{\mathcal{B}}$.

4.12 Lemma: (Vertical bundle associated to degenerate inertia tensor) Let $\mathcal{S} = (\mathcal{S}, \mathcal{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . If $\mathcal{V}(\text{Rgd}(\mathcal{B}; \mathcal{S})) = \ker(T\pi_{\mathcal{B}})$ is the vertical bundle for $\pi_{\mathcal{B}}$, then $\mathcal{V}(\text{Rgd}(\mathcal{B}; \mathcal{S})) = \ker(\mathbf{G}_{\mathcal{B}}^b)$.

Proof: First we note that

$$\begin{aligned} \ker(\mathbf{G}_{\mathcal{B}}^b)_{\Phi} &= \{(A, v) \mid mg(v, v) + \mathbf{G}(\mathbb{I}_c(\widetilde{R_{\Phi}^T \circ A}), \widetilde{R_{\Phi}^T \circ A}) = 0\} \\ &= \{(R_{\Phi} \circ \hat{\Omega}, 0) \mid \mathbf{G}(\mathbb{I}_c(\Omega), \Omega) = 0\} \\ &= \{(R_{\Phi} \circ \hat{\Omega}, 0) \mid \Omega \in \ker(\mathbb{I}_c)\}. \end{aligned}$$

Next let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, let $\Omega \in \mathfrak{i}(\mathcal{B}, \mu)$, and denote (with abuse of notation) by $e^{\hat{\Omega}\epsilon t}$ the element of $\text{Rgd}(\mathcal{B})$ consisting of a rotation by $e^{\hat{\Omega}\epsilon t}$ about X_c , i.e., the map

$$X \mapsto e^{\hat{\Omega}\epsilon t}(X - X_c).$$

Then compute

$$\begin{aligned} \mathbf{V}_\Phi(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})) &= \left\{ \frac{d}{dt} \Big|_{t=0} (\Phi \circ e^{\hat{\Omega}_\xi t}, r) \mid \xi \in \mathfrak{i}(\mathcal{B}, \mu) \right\} \\ &= \{(R_\Phi \circ \hat{\Omega}_\xi, 0) \mid \xi \in \mathfrak{i}(\mathcal{B}, \mu)\}, \end{aligned}$$

noting that the derivative of an affine map is its linear part. Combining the preceding two calculations gives the lemma. \blacksquare

Next we note that mere differentiation of the surjective submersion gives the surjective vector bundle mapping

$$T\pi_{\mathcal{B}}: \mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})/\mathfrak{l}(\mathcal{B})).$$

To give an alternative characterisation of this mapping, we first note that $\mathfrak{l}(\mathcal{B})$ -invariance of $\mathbf{G}_{\mathcal{B}}$ (by Corollary 4.7) means that $\mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))$ is $\mathfrak{l}(\mathcal{B})$ -invariant by Lemma 4.12. Thus we have an action of $\mathfrak{l}(\mathcal{B})$ on the quotient of $\mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))$ by $\mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))$ by

$$((\Phi, (A, v)) + \mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})), \Psi) \mapsto (\Phi \circ \Psi, (A \circ R_\Psi, v)) + \mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})), \quad (4.1)$$

keeping in mind Lemma 2.15. This gives us an alternative characterisation of $\mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})/\mathfrak{l}(\mathcal{B}))$.

4.13 Lemma: (The tangent bundle of the reduced configuration space) *Let $\mathcal{S} = (\mathcal{S}, \mathbf{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}, \mathbf{U}, \mathfrak{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Then the mapping*

$$\begin{aligned} \beta_{\mathcal{B}}: \mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})/\mathfrak{l}(\mathcal{B})) &\rightarrow (\mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))/\mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))/\mathfrak{l}(\mathcal{B})) \\ T\pi_{\mathcal{B}}(\Phi, (A, v)) &\mapsto [(\Phi, (A, v)) + \mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))] \end{aligned}$$

is an isomorphism of smooth vector bundles over the identity on $\mathrm{Rgd}(\mathcal{B}; \mathcal{S})/\mathfrak{l}(\mathcal{B})$.

Proof: First let us show that the mapping is well-defined. Suppose that $(\Phi_1, (A_1, v_1))$ and $(\Phi_2, (A_2, v_2))$ satisfy

$$T\pi_{\mathcal{B}}(\Phi_1, (A_1, v_1)) = T\pi_{\mathcal{B}}(\Phi_2, (A_2, v_2)).$$

Then, by Lemma 2.15, there exists $\Psi \in \mathfrak{l}(\mathcal{B})$ such that $\Phi_2 = \Phi_1 \circ \Psi$ and $A_2 = A_1 \circ R_\Psi$. By the definition (4.1) of the action of $\mathfrak{l}(\mathcal{B})$ on the quotient vector bundle, it follows that

$$\beta_{\mathcal{B}}([(\Phi_1, (A_1, v_1))]) = \beta_{\mathcal{B}}([(\Phi_2, (A_2, v_2))]),$$

and this gives the well-definedness of $\beta_{\mathcal{B}}$. It is clear from the definition that $\beta_{\mathcal{B}}$ preserves the fibres of the two vector bundles and is linear on each fibre. We can conclude that $\beta_{\mathcal{B}}$ is smooth by the following argument. The mapping

$$\begin{aligned} \bar{\beta}_{\mathcal{B}}: \mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S})) &\rightarrow (\mathbf{T}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))/\mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))/\mathfrak{l}(\mathcal{B})) \\ (\Phi, (A, v)) &\mapsto [(\Phi, (A, v)) + \mathbf{V}(\mathrm{Rgd}(\mathcal{B}; \mathcal{S}))] \end{aligned}$$

is smooth, being the composition of the quotient by a smooth subbundle followed by the quotient by a free and proper smooth group action. Now, to show that $\beta_{\mathcal{B}}$ is smooth, it suffices to show that $\beta_{\mathcal{B}}^* f$ is smooth for every smooth function f on

$(\mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))/\mathbb{V}(\text{Rgd}(\mathcal{B}; \mathcal{S}))/\mathbb{I}(\mathcal{B}))$. Let f be such a function and note that, since $T\pi_{\mathcal{B}}$ is a surjective submersion, one can easily show that $\beta_{\mathcal{B}}^* f$ is smooth if and only if $(T\pi_{\mathcal{B}})^* \beta_{\mathcal{B}}^* f$ is smooth. However,

$$(T\pi_{\mathcal{B}})^* \beta_{\mathcal{B}}^* f = \bar{\beta}_{\mathcal{B}}^* f,$$

and from this we conclude that $\beta_{\mathcal{B}}$ is indeed smooth. Now, by [Abraham, Marsden, and Ratiu 1988, Proposition 3.4.12], it follows that $\beta_{\mathcal{B}}$ is a smooth vector bundle map over the identity on $\text{Rgd}(\mathcal{B}; \mathcal{S})/\mathbb{I}(\mathcal{B})$. To show that $\beta_{\mathcal{B}}$ is injective, suppose that

$$\beta_{\mathcal{B}}(T\pi_{\mathcal{B}}((\Phi, (A, v)))) = 0.$$

Then $[(\Phi, (A, v)) + \mathbb{V}(\text{Rgd}(\mathcal{B}; \mathcal{S}))] = 0$ which means that

$$(\Phi, (A, v)) + \mathbb{V}(\text{Rgd}(\mathcal{B}; \mathcal{S})) = (\Phi \circ \Psi, (0 \circ R_{\Psi}, 0)) + \mathbb{V}(\text{Rgd}(\mathcal{B}; \mathcal{S})),$$

for some $\Psi \in \mathbb{I}(\mathcal{B})$, i.e., that $(A, v) \in \mathbb{V}_{\Phi}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$. Thus $(A, v) \in \ker(T_{\Phi}\pi_{\mathcal{B}})$, which gives injectivity of $\beta_{\mathcal{B}}$. Finally, to show that $\beta_{\mathcal{B}}$ is surjective, we need only note that every point in $(\mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))/\mathbb{V}(\text{Rgd}(\mathcal{B}; \mathcal{S}))/\mathbb{I}(\mathcal{B}))$, by definition, has the form $[(\Phi, (A, v)) + \mathbb{V}(\text{Rgd}(\mathcal{B}; \mathcal{S}))]$ for some $(\Phi, (A, v)) \in \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$. ■

Now we consider how the reduction of the set of configurations manifests itself in our velocity representations by the mappings τ_{spatial} and τ_{body} . We first give the lifted action of $\mathbb{I}(\mathcal{B})$ on $\mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$ under the spatial and body representations of velocity. To do so, we fix a point $x_0 \in \mathcal{S}$ and $X_c \in \mathcal{B}$ (which we need to do to define τ_{spatial} in any case) so that we have a diffeomorphism

$$\text{Rgd}(\mathcal{B}; \mathcal{S}) \simeq \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V}.$$

Since $\mathbb{I}(\mathcal{B})$ is a subgroup of $\text{SO}(\mathbb{U}, \mathbb{G}, \Theta) \times \{0\}$ under this diffeomorphism, the lifted action to the tangent bundle, under this representation, then has the form

$$\begin{aligned} \alpha_{\mathcal{B}}^T: \mathbb{T}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V}) \times \text{SO}(\mathbb{U}, \mathbb{G}, \Theta) &\rightarrow \mathbb{T}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V}) \\ ((R, r), (A, v), S) &\mapsto ((R \circ S, r), (A \circ S, 0)) \end{aligned}$$

by Lemma 2.15. We then compute

$$\begin{aligned} \tau_{\text{spatial}} \circ \alpha_{\mathcal{B}, S}^T \circ \tau_{\text{spatial}}^{-1}((R, r), (\omega, v)) &= \tau_{\text{spatial}} \circ \alpha_{\mathcal{B}, S}^T((R, r), (\hat{\omega} \circ R, v - r \times \omega)) \\ &= \tau_{\text{spatial}}((R \circ S, r), (\hat{\omega} \circ R \circ S, v - r \times \omega)) \\ &= ((R \circ S, r), (\omega, v)) \end{aligned} \quad (4.2)$$

and

$$\begin{aligned} \tau_{\text{body}} \circ \alpha_{\mathcal{B}, S}^T \circ \tau_{\text{body}}^{-1}((R, r), (\Omega, V)) &= \tau_{\text{body}} \circ \alpha_{\mathcal{B}, S}^T((R, r), (R \circ \hat{\Omega}, R(V))) \\ &= ((R \circ S, r), (R \circ \hat{\Omega} \circ S, R(V))) \\ &= ((R \circ S, r), ((S^T \circ \hat{\Omega} \circ S), S^T(V))) \\ &= ((R \circ S, r), (S^T(\Omega), S^T(V))), \end{aligned} \quad (4.3)$$

using Lemma 2.17. We denote these actions on

$$\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V} \times (\mathbb{V} \oplus \mathbb{V}), \quad \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V} \times (\mathbb{U} \oplus \mathbb{U})$$

by $\alpha_{\text{spatial}}^T$ and α_{body}^T , respectively.

We now introduce a rigid body (\mathcal{B}, μ) and define

$$\mathbf{Z}_{\text{spatial}} \subseteq \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V}), \quad \mathbf{Z}_{\text{body}} \subseteq \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U})$$

by

$$\mathbf{Z}_{\text{spatial}} = \{((R, r), (\omega, v)) \mid v = 0, \omega \in \ker(R \circ \mathbb{I}_c \circ R^T)\}$$

and

$$\mathbf{Z}_{\text{body}} = \{((R, r), (\Omega, V)) \mid v = 0, \Omega \in \ker(\mathbb{I}_c)\}.$$

At $(R, r) \in \text{Isom}^+(\mathcal{B}, \mu) \times \mathbf{V}$, the fibres of these subbundles are the degenerate spatial velocities and body velocities, respectively, for the body. These subbundles are $\mathfrak{l}(\mathcal{B})$ -invariant.

4.14 Lemma: (Invariance of subbundles of zero inertia) *The subbundles $\mathbf{Z}_{\text{spatial}}$ and \mathbf{Z}_{body} are invariant under the right-actions $\alpha_{\text{spatial}}^T$ and α_{body}^T of $\mathfrak{l}(\mathcal{B})$.*

Proof: This follows from the following two computations:

$$\begin{aligned} & \{\alpha_{\text{spatial}}^T(((R, r), (\omega, 0)), S) \mid \omega \in \ker(R \circ \mathbb{I}_c \circ R^T)\} \\ &= \{((R \circ S, r), (\omega, 0)) \mid \omega \in \ker(R \circ S \circ \mathbb{I}_c \circ S^T \circ R^T)\} \\ &= \{((R \circ S, r), (\omega, 0)) \mid \omega \in \ker((R \circ S) \circ \mathbb{I}_c \circ (R \circ S)^T)\} \end{aligned}$$

and

$$\begin{aligned} & \{\alpha_{\text{body}}^T(((R, r), (\Omega, 0)), S) \mid \Omega \in \ker(\mathbb{I}_c)\} \\ &= \{((R \circ S, r), (S^T(\Omega), 0)) \mid \Omega \in \ker(\mathbb{I}_c)\} \\ &= \{((R \circ S, r), (\Omega, 0)) \mid \Omega \in \ker(\mathbb{I}_c)\}, \end{aligned}$$

twice using Lemma 4.6. ■

Note, then, that the quotient bundles

$$(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V})) / \mathbf{Z}_{\text{spatial}}, \quad (\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U})) / \mathbf{Z}_{\text{body}}$$

inherit the right-actions $\alpha_{\text{spatial}}^T$ and α_{body}^T of $\mathfrak{l}(\mathcal{B})$ by (with an abuse of notation)

$$\alpha_{\text{spatial}}^T(((R, r), (\omega, v)) + \mathbf{Z}_{\text{spatial}}, S) = \alpha_{\text{spatial}}^T(((R, r), (\omega, v)), S) + \mathbf{Z}_{\text{spatial}}$$

and

$$\alpha_{\text{body}}^T(((R, r), (\Omega, V)) + \mathbf{Z}_{\text{body}}, S) = \alpha_{\text{body}}^T(((R, r), (\Omega, V)), S) + \mathbf{Z}_{\text{body}}.$$

We then have the following result that describes the spatial and body representations of velocity when the rigid body has degenerate inertia tensor.

4.15 Proposition: (Reduced rigid body velocities with degenerate inertia) *Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Then the mappings*

$$\begin{aligned} \tau_{\text{spatial}}^{\text{red}}: \mathbb{T}((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})/\mathbb{I}(\mathcal{B})) &\rightarrow ((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V}))/\mathbb{Z}_{\text{spatial}})/\mathbb{I}(\mathcal{B}) \\ T\pi_{\mathcal{B}}((R, r), (A, v)) &\mapsto [\tau_{\text{spatial}}((R, r), (A, v)) + \mathbb{Z}_{\text{spatial}}] \end{aligned}$$

and

$$\begin{aligned} \tau_{\text{body}}^{\text{red}}: \mathbb{T}((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})/\mathbb{I}(\mathcal{B})) &\rightarrow ((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U}))/\mathbb{Z}_{\text{body}})/\mathbb{I}(\mathcal{B}) \\ T\pi_{\mathcal{B}}((R, r), (A, v)) &\mapsto [\tau_{\text{body}}((R, r), (A, v)) + \mathbb{Z}_{\text{body}}] \end{aligned}$$

are vector bundle isomorphisms.

Proof: We start with the following data:

1. the vector bundles $\mathbb{T}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})$, $\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V})$, and $\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U})$, along with the vector bundle isomorphisms τ_{spatial} and τ_{body} ;
2. the right-actions $\alpha_{\mathcal{B}}^T$, $\alpha_{\text{spatial}}^T$, and α_{body}^T of $\mathbb{I}(\mathcal{B})$;
3. the Riemannian semimetric $\mathbf{G}_{\mathcal{B}}$ on $\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V}$ of Definition 3.7;
4. the fibre semimetric $\mathbf{G}_{\mathcal{B}}^{\text{spatial}}$ given in (3.2);
5. the fibre semimetric $\mathbf{G}_{\mathcal{B}}^{\text{body}}$ given in (3.3).

We then make the following observations:

6. the vector bundle isomorphisms τ_{spatial} and τ_{body} are equivariant (by the constructions (4.2) and (4.3));
7. the vector bundle isomorphisms τ_{spatial} and τ_{body} (and compositions of these) map the semimetrics above onto one another (by construction);
8. $\mathbf{V}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V}) = \ker(\mathbf{G}_{\mathcal{B}}^{\flat})$ (by Lemma 4.12);
9. $\mathbb{Z}_{\text{spatial}} = \ker((\mathbf{G}_{\mathcal{B}}^{\text{spatial}})^{\flat})$ and $\mathbb{Z}_{\text{body}} = \ker((\mathbf{G}_{\mathcal{B}}^{\text{body}})^{\flat})$ (by direct computation);
10. $\mathbb{T}((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})/\mathbb{I}(\mathcal{B}))$ is isomorphic as a vector bundle to the quotient by the right $\mathbb{I}(\mathcal{B})$ -action on the quotient vector bundle of $\mathbb{T}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})$ by the kernel of $\mathbf{G}_{\mathcal{B}}$, and the projection $T\pi_{\mathcal{B}}$ is just the natural projection of these two quotient operations (by Lemma 4.13).

An absorption of the preceding facts leads one to conclude that the three vector bundles

$$\begin{aligned} \mathbb{T}((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})/\mathbb{I}(\mathcal{B})), \quad &((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V}))/\mathbb{Z}_{\text{spatial}})/\mathbb{I}(\mathcal{B}), \\ &((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U}))/\mathbb{Z}_{\text{body}})/\mathbb{I}(\mathcal{B}) \end{aligned}$$

are obtained by starting with the three isomorphic vector bundles

$$\mathbb{T}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V}), \quad \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V}), \quad \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U}),$$

and then applying the construction of first quotienting by the kernel of a fibre semimetric and then quotienting by an action of $\mathbb{I}(\mathcal{B})$. Moreover, the three semimetrics are obtained from one another by transferring by the corresponding vector bundle isomorphisms, and the vector bundle isomorphisms are equivariant with respect to the group actions. \blacksquare

We comment that there is a fundamental difference in the mappings τ_{spatial} and τ_{body} , and $\tau_{\text{spatial}}^{\text{red}}$ and $\tau_{\text{body}}^{\text{red}}$, respectively.

4.16 Remark: (The nature of reduced rigid body velocities) If a rigid body has a non-degenerate inertia tensor, then $\tau_{\text{spatial}}^{\text{red}} = \tau_{\text{spatial}}$ and $\tau_{\text{body}}^{\text{red}} = \tau_{\text{body}}$. In this case, these vector bundle isomorphisms provide trivialisations of the tangent bundle of $\text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B})$:

$$\begin{aligned} ((R, r), (A, v)) &\mapsto ((R, r), (\widetilde{A \circ R^T}, v + r \times (\widetilde{A \circ R^T}))), \\ &((R, r), (A, v)) \mapsto ((R, r), (\widetilde{R^T \circ A}, R^T(v))). \end{aligned}$$

If the rigid body has zero inertia tensor, then, while the reduced velocity mappings $\tau_{\text{spatial}}^{\text{red}}$ and $\tau_{\text{body}}^{\text{red}}$ are no longer equal to their unreduced counterparts, they are still trivialisating as, upon quotienting by the internal symmetry group, they essentially take the form

$$(r, v) \mapsto (r, v), \quad (r, v) \mapsto (r, v).$$

However, when the inertia tensor is nonzero but degenerate, then the reduced velocity mappings are no longer trivialisating, as indeed the tangent bundle to the reduced configuration space is no longer trivialisable by Proposition 4.9(ii) and the Hairy Ball Theorem [Abraham, Marsden, and Ratiu 1988, Theorem 7.5.13]. This means that we do not have a convenient notational device in this case for separating out the spatial and body angular and translational velocities. •

4.4. Reduced kinetic energy metric. In this section we give what is the main point of our introduction of the reductions we have considered above, a Riemannian metric on $\text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B})$. It is this Riemannian metric that we shall use as our jumping off point for proving the equivalence of the Newton–Euler equations and the Euler–Lagrange equations in Section 8.

4.17 Lemma: (Riemannian metric associated with a degenerate inertia tensor) Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . The Riemannian semimetric $\mathbf{G}_{\mathcal{B}}$ on $\text{Rgd}(\mathcal{B}; \mathcal{S})$ descends to a Riemannian metric $\mathbf{G}_{0, \mathcal{B}}$ on $\text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B})$ that satisfies $\pi_{\mathcal{B}}^* \mathbf{G}_{0, \mathcal{B}} = \mathbf{G}_{\mathcal{B}}$, where $\pi_{\mathcal{B}}: \text{Rgd}(\mathcal{B}; \mathcal{S}) \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})/l(\mathcal{B})$ is the canonical projection.

Proof: We define $\mathbf{G}_{0, \mathcal{B}}$ by

$$\mathbf{G}_{0, \mathcal{B}}(T_{\Phi} \pi_{\mathcal{B}}(A_1, v_1), T_{\Phi} \pi_{\mathcal{B}}(A_2, v_2)) = \mathbf{G}_{\mathcal{B}}((A_1, v_1), (A_2, v_2))$$

for $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ and for $(A_1, v_1), (A_2, v_2) \in T_{\Phi}(\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V})$. By $l(\mathcal{B})$ -invariance of $\mathbf{G}_{\mathcal{B}}$ (Corollary 4.7), this expression for $\mathbf{G}_{0, \mathcal{B}}$ depends only on $\pi_{\mathcal{B}}(\Phi)$ (and not on Φ) and on $T_{\Phi} \pi_{\mathcal{B}}(A_a, v_a)$ (and not on (A_a, v_a)), $a \in \{1, 2\}$. Moreover, $\mathbf{G}_{0, \mathcal{B}}$ is nondegenerate since $\ker(T_{\Phi} \pi_{\mathcal{B}}) = \ker(\mathbf{G}_{\mathcal{B}}^{\flat})$ by Lemma 4.12. ■

The Riemannian metric $\mathbf{G}_{0, \mathcal{B}}$ we call the *kinetic energy metric* for the body.

Of course, one also has spatial and body versions of $\mathbf{G}_{0, \mathcal{B}}$ as fibre metrics on the vector bundles

$$((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{V} \oplus \mathbf{V}))/\mathbf{Z}_{\text{spatial}})/l(\mathcal{B}), \quad ((\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbf{V} \times (\mathbf{U} \oplus \mathbf{U}))/\mathbf{Z}_{\text{body}})/l(\mathcal{B})$$

that we denote by $\mathbf{G}_{0,\mathcal{B}}^{\text{spatial}}$ and $\mathbf{G}_{0,\mathcal{B}}^{\text{body}}$, respectively. Explicitly,

$$\begin{aligned} \mathbf{G}_{0,\mathcal{B}}^{\text{spatial}}(\tau_{\text{spatial}}^{\text{red}} \circ T_{\Phi}\pi_{\mathcal{B}}(\Phi, (A_1, v_1)), \tau_{\text{spatial}}^{\text{red}} \circ T_{\Phi}\pi_{\mathcal{B}}(\Phi, (A_2, v_2))) \\ = \mathbf{G}_{0,\mathcal{B}}(T_{\Phi}\pi_{\mathcal{B}}(A_1, v_1), T_{\Phi}\pi_{\mathcal{B}}(A_2, v_2)) \end{aligned}$$

and

$$\begin{aligned} \mathbf{G}_{0,\mathcal{B}}^{\text{body}}(\tau_{\text{body}}^{\text{red}} \circ T_{\Phi}\pi_{\mathcal{B}}(\Phi, (A_1, v_1)), \tau_{\text{body}}^{\text{red}} \circ T_{\Phi}\pi_{\mathcal{B}}(\Phi, (A_2, v_2))) \\ = \mathbf{G}_{0,\mathcal{B}}(T_{\Phi}\pi_{\mathcal{B}}(A_1, v_1), T_{\Phi}\pi_{\mathcal{B}}(A_2, v_2)). \end{aligned}$$

5. Interconnected rigid body systems: kinematics

Next we turn our attention to a general class of mechanical systems derived from interconnections of rigid bodies. What we do, essentially, is give a multibody version of the development of Sections 3 and 4, along with an introduction of constraints on configurations that arise from interconnections. In this section we also consider constraints on the velocities of the system. We consider throughout the effects of degenerate inertia tensors for the rigid bodies of the system, and much of the development is concerned with the technicalities required to do this.

5.1. Configuration manifold. We consider a Newtonian space model $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ and a finite collection $\mathcal{B}_a = (\mathcal{B}_a, \mathbf{U}_a, \mathbf{G}_a, \Theta_a)$, $a \in \{1, \dots, m\}$, of body reference spaces. In each of the body reference spaces \mathcal{B}_a , $a \in \{1, \dots, m\}$, we have a body (\mathcal{B}_a, μ_a) . The masses, centres of mass, and inertia tensors about the centre of mass of these bodies we denote by m_a , $X_{c,a}$, and $\mathbb{I}_{c,a}$, $a \in \{1, \dots, m\}$.

We first consider the artificial situation where all bodies may move freely in space, even allowing that multiple bodies occupy the same physical space.

5.1 Definition: (Free configuration manifold) Let \mathcal{S} be a Newtonian space model, let \mathcal{B}_a , $a \in \{1, \dots, m\}$, be body reference spaces, and let (\mathcal{B}_a, μ_a) , $a \in \{1, \dots, m\}$, be rigid bodies. The *free configuration manifold* for the system is $\mathbf{Q}_{\text{free}} = \prod_{a=1}^m \text{Rgd}(\mathcal{B}_a; \mathcal{S})$. •

We next wish to consider sets of admissible interconnections of the bodies $(\mathcal{B}_1, \dots, \mathcal{B}_m)$ in physical space.

5.2 Definition: (Physical configuration space, configuration space) Let \mathcal{S} be a Newtonian space model, let \mathcal{B}_a , $a \in \{1, \dots, m\}$, be body reference spaces, and let (\mathcal{B}_a, μ_a) , $a \in \{1, \dots, m\}$, be rigid bodies.

(i) A *physical configuration* of the bodies in \mathcal{S} is a subset

$$\Phi_1(\mathcal{B}_1) \times \dots \times \Phi_m(\mathcal{B}_m) \subseteq \prod_{a=1}^m \mathbf{S},$$

for some $\Phi_a \in \text{Rgd}(\mathcal{B}_a; \mathcal{S})$, $a \in \{1, \dots, m\}$.

(ii) A *physical configuration space* is a subset $\mathbf{P} \subseteq \prod_{a=1}^m \mathbf{2}^{\mathbf{S}}$ of physical configurations. A point in \mathbf{P} is called an *admissible physical configuration*.

- (iii) The *configuration space* associated with a physical configuration space P is the subset $Q \subseteq Q_{\text{free}}$ defined by

$$Q = \{(\Phi_1, \dots, \Phi_m) \mid \Phi_1(\mathcal{B}_1) \times \dots \times \Phi_m(\mathcal{B}_m) \in P\}.$$

A point in Q is called an *admissible configuration*. •

In general, the set of admissible configurations may be quite complicated. However, in this paper we shall consider the following situation, which is the one that is commonly (if tacitly) made in treatments at the level we are considering.

5.3 Definition: (Interconnected rigid body system) An *interconnected rigid body system* consists of the following data:

- (i) a Newtonian space model $\mathcal{S} = (S, V, g, \theta)$;
- (ii) body reference spaces $\mathcal{B}_a = (\mathcal{B}_a, U_a, G_a, \Theta_a)$ with rigid bodies (\mathcal{B}_a, μ_a) , $a \in \{1, \dots, m\}$;
- (iii) a physical configuration space P for which the associated configuration space Q is a smooth immersed submanifold of Q_{free} , called the *configuration manifold*, i.e., Q is the image of a smooth immersion $i: Q' \rightarrow Q_{\text{free}}$, and has the differentiable structure induced by this immersion. •

The preceding definition certainly captures the setting we wish to develop in this paper, by which we mean the “usual” framework for geometric mechanics. Indeed, it is quite general in this regard, since we allow for configuration manifolds that are merely immersed: generally one supposes that Q is embedded, but an example where it is immersed is given by Bullo and Lewis [2004, Exercise 4.2]. However, one can certainly do mechanics in frameworks more general than this, including what is normally referred to as “nonsmooth mechanics” where the set Q of admissible configurations may have boundaries or other nonsmooth phenomenon [Brogliato 1999, Kanno 2011].

5.4 Notation: (Configuration manifold)

1. We shall at times simply use words like “an interconnected rigid body system with configuration manifold Q ,” with the understanding that all of the other data needed to give meaning to these words is tacitly present.
2. Note that a point in Q is an element of Q_{free} . We shall thus denote a point in Q by $\Phi = (\Phi_1, \dots, \Phi_m)$. •

5.2. Reduced configuration manifold. We also wish to deal with situations when the rigid bodies have degenerate inertia tensors. As usual we have a Newtonian space model $\mathcal{S} = (S, V, g, \theta)$ and a finite collection $\mathcal{B}_a = (\mathcal{B}_a, U_a, G_a, \Theta_a)$, $a \in \{1, \dots, m\}$, of body reference spaces, with bodies $\mathcal{B}_1, \dots, \mathcal{B}_m$. We shall make the abbreviation $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$. We denote

$$I(\mathcal{B}) = I(\mathcal{B}_1) \times \dots \times I(\mathcal{B}_m),$$

and we note that $I(\mathcal{B})$ is a subgroup of $\prod_{a=1}^m \text{Rgd}(\mathcal{B}_a)$. This subgroup acts freely and properly on the right on Q_{free} with an action that we denote by $\alpha_{\mathcal{B}}$ and define by

$$\alpha_{\mathcal{B}}(\underbrace{(\Phi_1, \dots, \Phi_m)}_{\in Q_{\text{free}}}, \underbrace{(\Psi_1, \dots, \Psi_m)}_{\in I(\mathcal{B})}) = \underbrace{(\Phi_1 \circ \Psi_1, \dots, \Phi_m \circ \Psi_m)}_{\in Q_{\text{free}}}.$$

We call $l(\mathcal{B})$ the *internal symmetry group* of the bodies \mathcal{B} .

With the preceding constructions, we make the following definition.

5.5 Definition: (Reduced free configuration manifold) Let \mathcal{S} be a Newtonian space model, let \mathcal{B}_a , $a \in \{1, \dots, m\}$, be body reference spaces, and let (\mathcal{B}_a, μ_a) , $a \in \{1, \dots, m\}$, be rigid bodies. The *reduced free configuration manifold* for the system is $Q_{0, \text{free}} = Q_{\text{free}}/l(\mathcal{B})$. By $\pi_0: Q_{\text{free}} \rightarrow Q_{0, \text{free}}$ we denote the canonical projection. •

Of course, the reduced configuration manifold will be a product of manifolds as prescribed by Proposition 4.9, i.e., a product of components, with the a th component being diffeomorphic to one of (1) $\text{Isom}^+(\mathcal{B}; \mathcal{S}) \times V$ (when the $\mathbb{I}_{c,a}$ is nondegenerate), (2) $\mathbb{S}(V, g) \times V$ (when $\mathbb{I}_{c,a}$ is degenerate but nonzero), or (3) V (when $\mathbb{I}_{c,a}$ is zero).

Let us make a fundamental observation about the rôle of the internal symmetry group in the most general setting for configurations of rigid bodies.

5.6 Lemma: (The internal symmetry group acts on the configuration space) Let \mathcal{S} be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ be rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and let P be a physical configuration space with Q the corresponding configuration space. Then the subset $Q \subseteq Q_{\text{free}}$ is invariant under the right action of $l(\mathcal{B})$.

Proof: This follows since, for any $a \in \{1, \dots, m\}$, if $\Psi_a \in l(\mathcal{B}_a)$ then $\Psi_a(\mathcal{B}_a) = \mathcal{B}_a$, this by Lemma 4.5. ■

By the lemma, we can make the following definition.

5.7 Definition: (Reduced configuration space) Let \mathcal{S} be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ be rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and let P be a physical configuration space with Q the corresponding configuration space. The *reduced configuration space* is $Q_0 = Q/l(\mathcal{B})$. •

The following result relates this subgroup to our framework of configuration spaces.

5.8 Lemma: (Invariance of configuration space under internal symmetry) Let \mathcal{S} be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ be rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and let P be a physical configuration space with Q the corresponding configuration space. If the data describe an interconnected rigid body system, then Q_0 has the structure of a smooth manifold for which the projection $Q \mapsto Q_0$ defines for Q the structure of a principal $l(\mathcal{B})$ -bundle.

Proof: This follows since $l(\mathcal{B})$ is a closed Lie subgroup of $\prod_{a=1}^m \text{Rgd}(\mathcal{B}_a)$ [Abraham and Marsden 1978, Corollary 4.1.21], cf. the proof of Proposition 4.9. ■

Motivated by the lemma, we make the following definition.

5.9 Definition: (Reduced configuration manifold) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold Q . The *reduced configuration manifold* is $Q_0 = Q/l(\mathcal{B})$. By $\pi_0: Q \rightarrow Q_0$ we denote the canonical projection (accepting a slight abuse of notation). •

Note that the obvious equivariance of the inclusion of \mathbb{Q} in \mathbb{Q}_{free} gives the following commuting diagram:

$$\begin{array}{ccc} \mathbb{Q} & \hookrightarrow & \mathbb{Q}_{\text{free}} \\ \pi_0 \downarrow & & \downarrow \pi_0 \\ \mathbb{Q}_0 & \hookrightarrow & \mathbb{Q}_{0,\text{free}} \end{array} \quad (5.1)$$

That is to say, there is a natural inclusion of \mathbb{Q}_0 in $\mathbb{Q}_{0,\text{free}}$. This will be important to us in our development of the Euler–Lagrange equations from the Newton–Euler equations in Section 8.

5.3. Motion of interconnected rigid body systems. We next consider the velocity representations from Section 2.3 adapted to interconnected rigid body systems. We thus consider an interconnected rigid body system comprised of m rigid bodies and with configuration manifold \mathbb{Q} .

First we consider the free motions of the system, i.e., without imposing the requirement that motions remain in \mathbb{Q} . By taking the centre of mass for each body as an origin for its body reference space we establish, by a repeated application of the constructions of Corollary 2.14, a vector bundle monomorphism

$$\tau_c: \mathbb{T}\mathbb{Q}_{\text{free}} \rightarrow \mathbb{V}_c \triangleq \left(\prod_{a=1}^m \text{Rgd}(\mathcal{B}_a; \mathcal{S}) \right) \times \left(\bigoplus_{a=1}^m (\text{Hom}_{\mathbb{R}}(\mathbb{U}_a; \mathbb{V}) \oplus \mathbb{V}) \right)$$

defined by requiring that the restriction to $\mathbb{T}(\text{Rgd}(\mathcal{B}_a; \mathcal{S}))$ be the map $\tau_{X_{c,a}}$ of Corollary 2.14 for each $a \in \{1, \dots, m\}$.

In like manner, a repeated application of Lemma 2.13 and Definition 2.16 gives natural vector bundle isomorphisms

$$\tau_{\text{spatial}}: \mathbb{T}\mathbb{Q}_{\text{free}} \rightarrow \mathbb{V}_{\text{spatial}} \triangleq \left(\prod_{a=1}^m \text{Rgd}(\mathcal{B}_a; \mathcal{S}) \right) \times \left(\bigoplus_{a=1}^m \mathbb{V} \oplus \mathbb{V} \right)$$

and

$$\tau_{\text{body}}: \mathbb{T}\mathbb{Q}_{\text{free}} \rightarrow \mathbb{V}_{\text{body}} \triangleq \left(\prod_{a=1}^m \text{Rgd}(\mathcal{B}_a; \mathcal{S}) \right) \times \left(\bigoplus_{a=1}^m \mathbb{U}_a \oplus \mathbb{U}_a \right)$$

defined by requiring that the restriction to $\mathbb{T}(\text{Rgd}(\mathcal{B}_a; \mathcal{S}))$ be the maps $\tau_{\text{spatial},a}$ and $\tau_{\text{body},a}$ for each $a \in \{1, \dots, m\}$. Note that the definition of τ_{spatial} requires, not just a prescription of the centres of mass for all bodies as body origin, but also the choice of a spatial origin $x_0 \in \mathbb{S}$. However, this choice of spatial origin is not required for the construction of τ_{body} .

Finally, still considering free motions of our rigid bodies, we have reduced velocity mappings, which are vector bundle isomorphisms

$$\begin{aligned} \tau_{\text{spatial}}^{\text{red}}: \mathbb{T}\mathbb{Q}_{0,\text{free}} &\rightarrow \mathbb{V}_{\text{spatial}}^{\text{red}} \\ &\triangleq \left(\left(\prod_{a=1}^m \text{Isom}^+(\mathcal{B}_a; \mathcal{S}) \times \mathbb{V} \right) \times \left(\bigoplus_{a=1}^m \mathbb{V} \oplus \mathbb{V} \right) \right) / \tau_{\text{spatial}}(\mathbb{V}\mathbb{Q}_{\text{free}}) / \mathbb{I}(\mathcal{B}) \end{aligned}$$

and

$$\begin{aligned} \tau_{\text{body}}^{\text{red}} : \text{TQ}_{0,\text{free}} &\rightarrow \text{V}_{\text{body}}^{\text{red}} \\ &\triangleq \left(\left(\left(\prod_{a=1}^m \text{Isom}^+(\mathcal{B}_a; \mathcal{S}) \times \text{V} \right) \times \left(\bigoplus_{a=1}^m \text{U}_a \oplus \text{U}_a \right) \right) / \tau_{\text{body}}(\text{VQ}_{\text{free}}) \right) / \text{I}(\mathcal{B}), \end{aligned}$$

where $\text{VQ}_{\text{free}} = \ker(T\pi_0)$ is the vertical bundle for the free configuration space as a principal bundle over the reduced free configuration space.

Now it is an easy matter to include interconnections into the preceding development. Indeed, since $\text{Q} \subseteq \text{Q}_{\text{free}}$, we have $\text{TQ} \subseteq \text{TQ}_{\text{free}}$. We can thus restrict the vector bundle monomorphism τ_c to a vector bundle monomorphism with domain TQ and codomain V_c . We can also restrict the vector bundle isomorphisms τ_{spatial} and τ_{body} to vector bundle monomorphisms with domain TQ , the codomains being the restricted vector bundles $\text{V}_{\text{spatial}}|_{\text{Q}}$ and $\text{V}_{\text{body}}|_{\text{Q}}$, respectively. If the bodies (\mathcal{B}_a, μ_a) , $a \in \{1, \dots, m\}$, possess the internal symmetry group $\text{I}(\mathcal{B})$, then we have the inclusion $\text{TQ}_0 \subseteq \text{TQ}_{0,\text{free}}$. Thus the vector isomorphisms $\tau_{\text{spatial}}^{\text{red}}$ and $\tau_{\text{body}}^{\text{red}}$ restrict to vector bundle monomorphisms with domain TQ_0 , the codomains being the restricted vector bundles $\text{V}_{\text{spatial}}^{\text{red}}|_{\text{Q}_0}$ and $\text{V}_{\text{body}}^{\text{red}}|_{\text{Q}_0}$, respectively. Note also that the following diagram commutes:

$$\begin{array}{ccc} \text{TQ}_{\text{free}} & \xrightarrow{T\pi_0} & \text{TQ}_{0,\text{free}} \\ \pi_{\text{TQ}_{\text{free}}} \downarrow & & \downarrow \pi_{\text{TQ}_{0,\text{free}}} \\ \text{Q}_{\text{free}} & \xrightarrow{\pi_0} & \text{Q}_{0,\text{free}} \end{array}$$

While for TQ_{free} we have convenient trivialisations τ_{spatial} and τ_{body} , we do not generally have such trivialisations for the reduced velocities, cf. Remark 4.16.

5.10 Notation: (Tangent bundle of configuration manifold) Given the above representation of TQ_{free} via the vector bundle mapping τ_c , we shall very often write an element of TQ as

$$((\Phi_1, \dots, \Phi_m), ((A_1, v_1), \dots, (A_m, v_m))).$$

When additional brevity is needed, we may sometimes write such a point simply as $(\Phi, (\mathbf{A}, \mathbf{v}))$, accepting a mild abuse of notation in this case. •

Now let us consider motions of systems.

5.11 Definition: (Motion, reduced motion) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold Q . Then:

- (i) a *motion* for the system is a curve $\phi: \mathbb{T}' \rightarrow \text{Q}$ whose domain is a time interval in a time axis \mathbb{T} ;
- (ii) the *reduced motion* associated to a motion $\phi: \mathbb{T}' \rightarrow \text{Q}$ is the curve $\phi_0 \triangleq \pi_0 \circ \phi: \mathbb{T}' \rightarrow \text{Q}_0$. •

For a differentiable motion $\phi: \mathbb{T}' \rightarrow \text{Q}$, we can define corresponding spatial velocities $\omega_{\phi_a}: \mathbb{T}' \rightarrow \text{V}$ and $v_{\phi_a}: \mathbb{T}' \rightarrow \text{V}$, $a \in \{1, \dots, m\}$, by requiring that

$$\begin{aligned} \tau_{\text{spatial}} \circ \phi'(t) &= ((\phi_1(t), \dots, \phi_m(t)), \\ &((\omega_{\phi_1}(t), v_{\phi_1}(t)), \dots, (\omega_{\phi_m}(t), v_{\phi_m}(t)))), \quad t \in \mathbb{T}', \end{aligned}$$

and body velocities $\Omega_{\phi_a}: \mathbb{T}' \rightarrow \mathbf{U}_a$ and $V_{\phi_a}: \mathbb{T}' \rightarrow \mathbf{U}_a$, $a \in \{1, \dots, m\}$, by requiring that

$$\begin{aligned} \tau_{\text{body}} \circ \phi'(t) &= ((\phi_1(t), \dots, \phi_m(t)), \\ &\quad ((\Omega_{\phi_1}(t), V_{\phi_1}(t)), \dots, (\Omega_{\phi_m}(t), V_{\phi_m}(t))))), \quad t \in \mathbb{T}', \end{aligned}$$

using the notation of Definition 2.16. One also has velocities for the reduced motion, and spatial and body representations of these. These are defined by

$$\phi'_0(t) = T_{\phi(t)}\pi_0 \circ \phi'(t), \quad \tau_{\text{spatial}}^{\text{red}} \circ \phi'_0(t), \quad \tau_{\text{body}}^{\text{red}} \circ \phi'_0(t).$$

Since these have no particularly convenient representation in general, we do not attempt to give expressions for these.

5.4. Riemannian semimetrics and metrics for interconnected rigid body systems. It is a relatively straightforward matter to adapt the constructions of Sections 3.3 and 4.4 to the situation where we have interconnections. In this section we perform the more or less obvious constructions and give the resulting notation we shall use. We shall make a few statements in this multibody setting that we shall not prove, in all cases these following easily from single body analogues.

We let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B}_a = (\mathbf{B}_a, \mathbf{U}_a, \mathbf{G}_a, \Theta_a)$, $a \in \{1, \dots, m\}$, be body reference spaces with bodies (\mathcal{B}_a, μ_a) , $a \in \{1, \dots, m\}$. By applying the construction of Definition 3.7 to each component of the product \mathbf{Q}_{free} , we obtain a Riemannian semimetric on \mathbf{Q}_{free} that we denote by $\mathbf{G}_{\mathcal{B}}$. Explicitly,

$$\mathbf{G}_{\mathcal{B}}((\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2)) = \sum_{a=1}^m \mathbf{G}_{\mathcal{B}_a}((A_{1,a}, v_{1,a}), (A_{2,a}, v_{2,a})), \quad (5.2)$$

for $(\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2) \in \mathbf{T}_{\Phi} \mathbf{Q}_{\text{free}}$ and $\Phi \in \mathbf{Q}_{\text{free}}$. This has spatial and body representations—by applying the constructions of (3.2) and (3.3) componentwise—as fibre semimetrics on the vector bundles $\mathbf{V}_{\text{spatial}}$ and \mathbf{V}_{body} . These fibre metrics we denote by $\mathbf{G}_{\mathcal{B}}^{\text{spatial}}$ and $\mathbf{G}_{\mathcal{B}}^{\text{body}}$, respectively. Explicitly,

$$\mathbf{G}_{\mathcal{B}}^{\text{spatial}}(\tau_{\text{spatial}}(\mathbf{A}_1, \mathbf{v}_1), \tau_{\text{spatial}}(\mathbf{A}_2, \mathbf{v}_2)) = \mathbf{G}_{\mathcal{B}}((\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2)),$$

and

$$\mathbf{G}_{\mathcal{B}}^{\text{body}}(\tau_{\text{body}}(\mathbf{A}_1, \mathbf{v}_1), \tau_{\text{body}}(\mathbf{A}_2, \mathbf{v}_2)) = \mathbf{G}_{\mathcal{B}}((\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2)),$$

for $(\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2) \in \mathbf{T}_{\Phi} \mathbf{Q}_{\text{free}}$ and $\Phi \in \mathbf{Q}_{\text{free}}$. By Corollary 4.7, $\mathbf{G}_{\mathcal{B}}$ is invariant under the right action of $l(\mathcal{B})$ on \mathbf{Q}_{free} . Thus, just as in Lemma 4.17, the Riemannian semimetric $\mathbf{G}_{\mathcal{B}}$ descends to a Riemannian metric on $\mathbf{Q}_{0, \text{free}}$. We denote this Riemannian metric by $\mathbf{G}_{0, \mathcal{B}}$. Explicitly,

$$\mathbf{G}_{0, \mathcal{B}}(T_{\Phi}\pi_0(\mathbf{A}_1, \mathbf{v}_1), T_{\Phi}\pi_0(\mathbf{A}_2, \mathbf{v}_2)) = \mathbf{G}_{\mathcal{B}}((\mathbf{A}_1, \mathbf{v}_2), (\mathbf{A}_2, \mathbf{v}_2)), \quad (5.3)$$

for $(\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2) \in \mathbf{T}_{\Phi} \mathbf{Q}_{\text{free}}$ and $\Phi \in \mathbf{Q}_{\text{free}}$. This has spatial and body representations as a fibre metric on the vector bundles $\mathbf{V}_{\text{spatial}}^{\text{red}}$ and $\mathbf{V}_{\text{body}}^{\text{red}}$ that we denote by $\mathbf{G}_{0, \mathcal{B}}^{\text{spatial}}$ and $\mathbf{G}_{0, \mathcal{B}}^{\text{body}}$, respectively. Explicitly,

$$\mathbf{G}_{0, \mathcal{B}}^{\text{spatial}}(\tau_{\text{spatial}}^{\text{red}} \circ T_{\Phi}\pi_0(\mathbf{A}_1, \mathbf{v}_1), \tau_{\text{spatial}}^{\text{red}} \circ T_{\Phi}\pi_0(\mathbf{A}_2, \mathbf{v}_2)) = \mathbf{G}_{\mathcal{B}}((\mathbf{A}_1, \mathbf{v}_2), (\mathbf{A}_2, \mathbf{v}_2))$$

and

$$\mathbf{G}_{0,\mathcal{B}}^{\text{body}}(\tau_{\text{body}}^{\text{red}} \circ T_{\Phi}\pi_0(\mathbf{A}_1, \mathbf{v}_1), \tau_{\text{body}}^{\text{red}} \circ T_{\Phi}\pi_0(\mathbf{A}_2, \mathbf{v}_2)) = \mathbf{G}_{\mathcal{B}}((\mathbf{A}_1, \mathbf{v}_2), (\mathbf{A}_2, \mathbf{v}_2)),$$

for $(\mathbf{A}_1, \mathbf{v}_1), (\mathbf{A}_2, \mathbf{v}_2) \in \mathbf{T}_{\Phi}(\mathbf{Q}_{\text{free}})$ and $\Phi \in \mathbf{Q}_{\text{free}}$.

Now we consider the effects of interconnections. Thus we suppose that, along with the above data, we have a set of admissible configurations defining an interconnected rigid body system with configuration manifold \mathbf{Q} . Simply by restriction, the Riemannian semimetric $\mathbf{G}_{\mathcal{B}}$ induces a Riemannian semimetric on \mathbf{Q} . We shall denote this Riemannian semimetric by \mathbf{G} . Since $\mathbf{G}_{\mathcal{B}}$ and \mathbf{Q} are $l(\mathcal{B})$ -invariant, the Riemannian semimetric \mathbf{G} on \mathbf{Q} descends to a Riemannian metric \mathbf{G}_0 on \mathbf{Q}_0 . This Riemannian metric agrees with the restriction of the Riemannian metric on $\mathbf{Q}_{0,\text{free}}$ to \mathbf{Q}_0 by virtue of the inclusion defined by the commuting diagram (5.1).

5.5. Velocity constraints for interconnected rigid body systems. We will consider velocity constraints for systems of interconnected rigid bodies that model physical phenomenon like rolling contact. We do so by first introducing the notion of a constraint on the motion of a single rigid body. This will allow us to give a physically meaningful construction of velocity constraints by amalgamating such constraints over the various bodies of an interconnected rigid body system.

Thus we start with the notion of a velocity constraint for a rigid body.

5.12 Definition: (Primary velocity constraint, constraint subspace) Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ and $X_0 \in \mathcal{B}$.

- (i) A **primary velocity constraint** for \mathcal{B} at (Φ, X_0) is a subspace $\mathbf{C}_{\Phi, X_0} \subseteq \mathbf{V} \oplus \mathbf{V}$ such that

$$\{(\omega, 0) \in \mathbf{V} \oplus \mathbf{V} \mid \omega \in \ker(R_{\Phi} \circ \mathbb{I}_c \circ R_{\Phi}^T)\} \subseteq \mathbf{C}_{\Phi, X_0}. \quad (5.4)$$

- (ii) The **constraint subspace** associated to a primary velocity constraint \mathbf{C}_{Φ, X_0} is the subspace $\mathbf{D}_{\Phi, X_0} \subseteq \mathbf{T}_{\Phi}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$ given by

$$\mathbf{D}_{\Phi, X_0} = \{(\Phi, (A, v)) \mid (A \circ R_{\Phi}^T, v + A(X_0 - X_c)) \in \mathbf{C}_{\Phi, X_0}\},$$

using the identification of Corollary 2.14 with respect to X_c . •

Let us give the physical meaning of the preceding mathematical definitions. First, let us consider a primary velocity constraint \mathbf{C}_{Φ, X_0} and let $(\omega, v) \in \mathbf{C}_{\Phi, X_0}$. We should think of v as being a possible spatial translational velocity of the point $\Phi(X_0)$ and ω as being a possible spatial angular velocity of the body $\Phi(\mathcal{B})$ about the point $\Phi(X_0)$. Thus a primary velocity constraint is a constraint on the velocities of the body about the reference point $\Phi(X_0)$. The condition (5.4) means that we should not constrain the inertialess motions of the body. Now let us understand the meaning of the constraint subspace associated with a primary velocity constraint. Let us suppose that we have a differentiable rigid motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ such that $\phi(t_0) = \Phi$. Then the spatial motion of the point X_0 in the body is given by

$$t \mapsto \phi(t)(X_0) = \phi(t)(X_c) + R_{\phi}(t)(X_0 - X_c).$$

Differentiating this expression at t_0 gives

$$\left. \frac{d}{dt} \right|_{t=t_0} \phi(t)(X_0) = \left. \frac{d}{dt} \right|_{t=t_0} \phi(t)(X_c) + \omega_\phi(t) \times (R_\phi(t)(X_0 - X_c)).$$

If we require that

$$\left(\omega_\phi(t), \left. \frac{d}{dt} \right|_{t=t_0} \phi(t)(X_0) \right) \in \mathbb{C}_{\Phi, X_0},$$

then

$$\left(\dot{R}_\phi(t), \left. \frac{d}{dt} \right|_{t=t_0} \phi(t)(X_c) \right) \in \mathbb{D}_{\Phi, X_0},$$

using the identification of $\mathbb{T}_\Phi(\text{Rgd}(\mathcal{B}; \mathcal{S}))$ from Corollary 2.14. The preceding definition gives as a precise way to transfer the physical data of how one might typically prescribe a primary velocity constraint into a subspace tangent to the configuration space of the rigid body.

Now we present the way in which one makes these definitions for interconnections of rigid bodies.

5.13 Definition: (Physical velocity constraint, constraint subspace, constraint distribution) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathbb{Q} .

- (i) A *physical velocity constraint* is an assignment, to each $\Phi = (\Phi_1, \dots, \Phi_m) \in \mathbb{Q}$, the following data:
 - (a) for each $a \in \{1, \dots, m\}$, $k_{\Phi, a} \in \mathbb{Z}_{\geq 0}$ satisfying $k_{\alpha_{\mathcal{B}, \Psi}(\Phi), a} = k_{\Phi, a}$ for every $\Psi \in \mathbb{I}(\mathcal{B})$;
 - (b) for each $a \in \{1, \dots, m\}$ and $j \in \{1, \dots, k_{\Phi, a}\}$, a point $X_{\Phi, a, j} \in \mathcal{B}_a$ satisfying $X_{\alpha_{\mathcal{B}, \Psi}(\Phi), a, j} = X_{\Phi, a, j}$ for every $\Psi \in \mathbb{I}(\mathcal{B})$;
 - (c) for each $a \in \{1, \dots, m\}$ and $j \in \{1, \dots, k_{\Phi, a}\}$, a primary velocity constraint $\mathbb{C}_{\Phi, a, j}$ for \mathcal{B}_a at $(\Phi_a, X_{\Phi, a, j})$ satisfying $\mathbb{C}_{\alpha_{\mathcal{B}, \Psi}(\Phi), a, j} = \mathbb{C}_{\Phi, a, j}$ for every $\Psi \in \mathbb{I}(\mathcal{B})$.
- (ii) For $\Phi \in \mathbb{Q}$, $a \in \{1, \dots, m\}$, and $j \in \{1, \dots, k_{\Phi, a}\}$, the *constraint subspace* is the subspace $\mathbb{D}_{\Phi, a, j} \subseteq \mathbb{T}_\Phi \mathbb{Q}_{\text{free}}$ given by

$$\mathbb{D}_{\Phi, a, j} = \{(\Phi, (A, v)) \in \mathbb{T}_\Phi \mathbb{Q}_{\text{free}} \mid (A_a \circ R_{\Phi_a}^T, v_a + A_a(X_{\Phi, a, j} - X_{c, a})) \in \mathbb{C}_{\Phi, a, j}\}.$$

- (iii) The *constraint distribution* associated to a physical velocity constraint is the assignment, for each $\Phi \in \mathbb{Q}$, of the subspace $\mathbb{D}_\Phi \subseteq \mathbb{T}_\Phi \mathbb{Q}$ given by

$$\mathbb{D}_\Phi = \mathbb{T}_\Phi \mathbb{Q} \cap \left(\bigcap_{a=1}^m \bigcap_{j=1}^{k_{\Phi, a}} \mathbb{D}_{\Phi, a, j} \right). \quad \bullet$$

The physical meaning of this mathematical definition is the natural adaptation of that above for single rigid bodies, allowing for the facts that (1) each body may have primary velocity constraints applied at multiple points and (2) the allowable velocities for the a th body might depend on the configurations of the other bodies. The constraint distribution is the set of velocities satisfying *all* constraints, as well as the constraint of remaining in \mathbb{Q} .

The following basic attribute of constraints will allow us to reduce these by the internal symmetry group.

5.14 Lemma: (Invariance of constraint distributions under internal symmetry)

Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathcal{Q} . If we are given physical velocity constraints giving rise to a constraint distribution $\mathcal{D} \subseteq \mathcal{T}\mathcal{Q}$, then \mathcal{D} is invariant under the right-action of $\mathcal{I}(\mathcal{B})$ in the sense that $T_{\Phi}\alpha_{\mathcal{B},\Psi}(\mathcal{D}_{\Phi}) = \mathcal{D}_{\alpha_{\mathcal{B},\Psi}(\Phi)}$ for every $\Phi \in \mathcal{Q}$ and $\Psi \in \mathcal{I}(\mathcal{B})$, where $\alpha_{\mathcal{B}}$ denotes the right-action of $\mathcal{I}(\mathcal{B})$.

Proof: Since $\mathcal{I}(\mathcal{B})$ acts by rotations about the centres of mass of the bodies, we have

$$T_{\Phi}\alpha_{\mathcal{B},\Psi}((A_1, v_1), \dots, (A_m, v_m)) = ((A_1 \circ R_{\Psi_1}, v_1), \dots, (A_m \circ R_{\Psi_m}, v_m))$$

by Lemma 2.15. Thus the result follows from (1) the following computations:

$$(A_a \circ R_{\Psi_a}) \circ (R_{\Phi_a} \circ R_{\Psi_a})^T = A_a \circ R_{\Phi_a}^T, \quad a \in \{1, \dots, m\},$$

and

$$\begin{aligned} v_a + A_a \circ R_{\Psi_a}(X_{\alpha_{\mathcal{B},\Psi}(\Phi),a,j} - X_{c,a}) &= v_a + A_a \circ R_{\Psi_a}(X_{\Phi,a,j} - X_{c,a}) \\ &= v_a + A_a(X_{\Phi,a,j} - X_{c,a}), \end{aligned}$$

$a \in \{1, \dots, m\}$, $j \in \{1, \dots, k_{\Phi,a}\}$, whenever

$$((A_1, v_1), \dots, (A_m, v_m)) \in \mathcal{D}_{\Phi},$$

along with (2) the fact that $\mathcal{C}_{\alpha_{\mathcal{B},\Psi}(\Phi),a,j} = \mathcal{C}_{\Phi,a,j}$. In the second of the above computations we have used Lemma 4.5. \blacksquare

Generally speaking, of course, a constraint distribution as prescribed above will have no nice properties; indeed, one must ask that all subspaces “fit together” in a nice way. This is done according to the following definition.

5.15 Definition: (Attributes of constraint distributions) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathcal{Q} . Suppose that we have physical velocity constraints giving rise to the constraint distribution \mathcal{D} .

- (i) The velocity constraint is *cosmooth* if, for every $\Phi_0 \in \mathcal{Q}$, there exists a neighbourhood \mathcal{U} of Φ_0 and smooth one-forms $(\alpha_a)_{a \in A}$ on \mathcal{U} such that

$$\text{ann}(\mathcal{D}_{\Phi}) = \text{span}(\alpha_a(\Phi) \mid a \in A)$$

for every $\Phi \in \mathcal{U}$.

- (ii) The velocity constraint has *locally constant rank* if it is cosmooth and the function

$$\begin{aligned} \text{rank}_{\mathcal{D}}: \mathcal{Q} &\rightarrow \mathbb{Z}_{\geq 0} \\ \Phi &\mapsto \dim(\mathcal{D}_{\Phi}) \end{aligned}$$

is locally constant. \bullet

It is valid to speculate on why we ask that the annihilating codistribution for a constraint distribution be smooth, and not the constraint distribution itself. The reason for this is that, in practice, it is the constraint *forces* that vary smoothly, not the spaces of admissible velocities. We shall look at constraint forces in Section 6.7. In this paper we shall only consider locally constant rank velocity constraints, and in this case cosmoothness of the velocity constraint is equivalent to smoothness of the constraint distribution itself. As far as we are aware, there does not exist a satisfactory physical theory for velocity constraints that do not have locally constant rank, although there do exist physical systems with constraints of this type.

Note that there are many interesting attributes of constraint distributions that can be discussed using differential geometric techniques, e.g., integrability in the sense of Fröbenius and the Chow–Rashevsky Theorem. These, however, belong more properly to the subject of geometric mechanics itself, and not as much to the physical foundations such as interest us here. We refer to books such as [Bloch 2003] and [Bullo and Lewis 2004, §4.5] for details on these and other topics concerning “nonholonomic constraints” in geometric mechanics.

Finally, let us indicate how constraint distributions descend to distributions on the reduced configuration space.

5.16 Lemma: (Constraint distributions descend to the reduced configuration space) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathcal{Q} . Suppose that we have physical velocity constraints giving rise to the constraint distribution D . If we denote*

$$D_{0,\pi_0(\Phi)} = T_{\Phi}\pi_0(D_{\Phi}),$$

then the following statements hold:

- (i) D_0 is well-defined, i.e., it depends only on $\pi_0(\Phi)$ and not on Φ ;
- (ii) if D is cosmooth, then so is D_0 ;
- (iii) if D has locally constant rank, then so does D_0 .

Proof: (i) This follows at once from $l(\mathcal{B})$ -invariance of D , along with the fact that $\pi_0 = \pi_0 \circ \alpha_{\mathcal{B},\Psi}$ for every $\Psi \in l(\mathcal{B})$.

(ii) Since $\pi_0: \mathcal{Q} \rightarrow \mathcal{Q}_0$ is a principal $l(\mathcal{B})$ -bundle, we may place on it a smooth connection [Kobayashi and Nomizu 1963, Theorem II.2.1]. This gives a decomposition $T\mathcal{Q} = V\mathcal{Q} \oplus H\mathcal{Q}$ for some complement $H\mathcal{Q}$ of $V\mathcal{Q} \triangleq \ker(T\pi_0)$. We then have a corresponding decomposition $T^*\mathcal{Q} = V^*\mathcal{Q} \oplus H^*\mathcal{Q}$ of the cotangent bundle. For $\Phi \in \mathcal{Q}$ and $\Phi_0 = \pi_0(\Phi)$, we have isomorphisms

$$T_{\Phi}\pi_0|_{H_{\Phi}\mathcal{Q}}: H_{\Phi}\mathcal{Q} \rightarrow T_{\Phi_0}\mathcal{Q}_0, \quad (T_{\Phi}\pi_0|_{H_{\Phi}\mathcal{Q}})^*: T_{\Phi_0}^*\mathcal{Q}_0 \rightarrow H_{\Phi}^*\mathcal{Q}.$$

Let us abbreviate the second of these maps by i_{Φ} .

Now let $\Phi_0 \in \mathcal{Q}_0$, let $\mathcal{U} \subseteq \mathcal{Q}_0$ be a neighbourhood of Φ_0 , and let $\sigma: \mathcal{U} \rightarrow \mathcal{Q}$ be a local section of $\pi_0: \mathcal{Q} \rightarrow \mathcal{Q}_0$. Let $\Phi = \sigma(\Phi_0)$, let $\mathcal{V} \subseteq \mathcal{Q}$ be a neighbourhood of Φ , and let $(\alpha_a)_{a \in A}$ be one-forms on \mathcal{V} such that

$$D_{\Phi'} = \text{span}(\alpha_a(\Phi') \mid a \in A)$$

for each $\Phi' \in \mathcal{V}$. Suppose that \mathcal{U} is sufficiently small that $\sigma(\mathcal{U}) \subseteq \mathcal{V}$. For $a \in A$, let α_a^h be the projection of α_a onto $\mathbf{H}^*\mathbf{Q}$, noting that α_a^h is smooth. We then define a smooth one-form β_a on \mathcal{U} by

$$\beta_a(\Phi'_0) = i_{\sigma(\Phi'_0)}^{-1} \circ \alpha_a^h(\sigma(\Phi'_0)), \quad \Phi'_0 \in \mathcal{U}.$$

A bit of linear algebra then shows that

$$D_{0, \Phi'_0} = \text{span}(\beta_a(\Phi'_0) \mid a \in A),$$

which gives this part of the lemma.

(iii) This follows from the following observation about dimensions:

$$\dim(D_\Phi) = \dim(\mathbf{l}(\mathcal{B})) + \dim(D_{0, \pi_0(\Phi)}),$$

this because of Lemma 5.14 and the fact that $\mathbf{VQ}_{\text{free}}|\mathbf{Q} \subseteq \mathbf{D}$ by definition of primary velocity constraint. \blacksquare

The preceding lemma allows us to make the following definition.

5.17 Definition: (Reduced constraint distribution) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathbf{Q} . Suppose that we have physical velocity constraints giving rise to the constraint distribution \mathbf{D} . The distribution $D_0 \subseteq \mathbf{TQ}_0$ is the *reduced constraint distribution*. \bullet

Since the preceding constructions may not be entirely transparent, we illustrate them with a simple example.

5.18 Example: (Sliding rod) The physical system we consider is a rod, one tip of which slides on a plane. We shall be very concrete, and so we let $\mathbf{S} = \mathbf{B} = \mathbb{R}^3$ which have the standard structure as affine spaces modelled on the vector spaces $\mathbf{V} = \mathbf{U} = \mathbb{R}^3$. We use the standard inner products and orientations on both \mathbf{U} and \mathbf{V} . We take the spatial origin to be $\mathbf{x}_0 = \mathbf{0}$ and use the standard orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ for \mathbb{R}^3 . The body in this case we define as

$$\mathcal{B} = \{(0, 0, z) \in \mathbb{R}^3 \mid z \in [a, b]\},$$

i.e., the body is an infinitely thin rod of length $b - a$. By Lemma 3.4 we have $X_c = (0, 0, z_c)$ for some $z_c \in (a, b)$. We assume, without loss of generality, that $z_c = 0$. In \mathbf{S} we define a plane

$$\mathbf{P} = \{(x, y, 0) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}.$$

The set of physical configurations we take to be

$$\mathbf{P} = \{\Phi(\mathcal{B}) \mid \Phi(0, 0, a) \in \mathbf{P}\},$$

i.e., the tip of the rod at $(0, 0, a)$ is required to reside in the plane \mathbf{P} as in Figure 2. The velocity constraint is that the motion of the rod be such that the velocity of the point of contact of the rod with \mathbf{P} be tangent to \mathbf{P} . Thus, according to Definition 5.12, we take $\mathbf{X}_0 = (0, 0, a)$ and let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ be an admissible configuration, i.e., one for which $\Phi(\mathbf{X}_0) \in \mathbf{P}$. The translational velocities of the point $\Phi(\mathbf{X}_0)$ allowed by the velocity

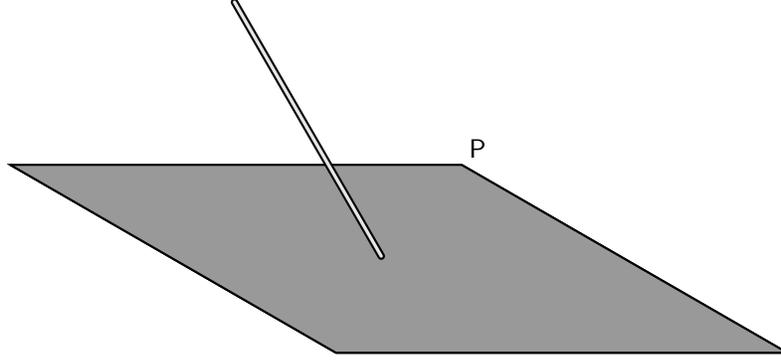


Figure 2. Rod with tip constrained to move in a plane

constraint are then of the form $(v_x, v_y, 0) \in \mathcal{V}$. A motion of the system has the form $t \mapsto \phi(t) = (\mathbf{R}(t), \mathbf{r}(t)) \in \text{SO}(3) \times \mathbb{R}^3$, where $\mathbf{r}(t) = \phi(t)(\mathbf{X}_c)$. Thus

$$\phi(t)(\mathbf{X}_0) = \mathbf{r}(t) + \mathbf{R}(t)(\mathbf{X}_0) \quad \implies \quad \frac{d}{dt}\phi(t)(\mathbf{X}_0) = \dot{\mathbf{r}}(t) + \boldsymbol{\omega}(t) \times (\mathbf{R}(t)(\mathbf{X}_0)).$$

Thus the primary velocity constraint for the system at (Φ, \mathbf{X}_0) is

$$\mathcal{C}_{\Phi, \mathbf{X}_0} = \{(\boldsymbol{\omega}, \mathbf{v}) \in \mathbb{R}^3 \oplus \mathbb{R}^3 \mid \mathfrak{g}(\mathbf{e}_3, \mathbf{v} + \boldsymbol{\omega} \times (\mathbf{R}_{\Phi}(\mathbf{X}_0))) = 0\}.$$

To see that this is indeed a primary velocity constraint, note that $\ker(\mathbb{I}_c) = \text{span}(\mathbf{e}_3)$. Therefore,

$$\boldsymbol{\omega} \in \ker(\mathbf{R}_{\Phi} \circ \mathbb{I}_c \circ \mathbf{R}_{\Phi}^T) \quad \iff \quad \boldsymbol{\omega} = a\mathbf{R}_{\Phi}(\mathbf{e}_3)$$

for some $a \in \mathbb{R}$. Thus, if $\boldsymbol{\omega} \in \ker(\mathbf{R}_{\Phi} \circ \mathbb{I}_c \circ \mathbf{R}_{\Phi}^T)$, then

$$\mathfrak{g}(\mathbf{e}_3, \mathbf{0} + \boldsymbol{\omega} \times (\mathbf{R}_{\Phi}(\mathbf{X}_0))) = 0$$

since $\boldsymbol{\omega}$ and $\mathbf{R}_{\Phi}(\mathbf{X}_0)$ are both collinear with $\mathbf{R}_{\Phi}(\mathbf{e}_3)$. The corresponding constraint subspace is

$$\mathcal{D}_{\Phi, \mathbf{X}_0} = \{(\Phi, (\mathbf{A}, \mathbf{v})) \in \text{Rgd}(\mathcal{B}; \mathcal{S}) \times (\mathbb{R}^{3 \times 3} \oplus \mathbb{R}^3) \mid \mathfrak{g}(\mathbf{e}_3, \mathbf{v} + \mathbf{A}(\mathbf{X}_0)) = 0\}.$$

We note that, because of the simplicity of this example, the tangent bundle of \mathcal{Q} at Φ is exactly the constraint subspace $\mathcal{D}_{\Phi, \mathbf{X}_0}$. Thus the constraint distribution is defined by $\mathcal{D}_{\Phi} = \mathcal{D}_{\Phi, \mathbf{X}_0}$. •

6. Interconnected rigid body systems: forces and torques

The last mechanical element in our formulation is that of forces and torques. As usual, we present these starting from an elemental physical formulation, giving the mathematical meaning of all basic notions. Part of the formulation necessarily includes a discussion of forces and torques required to maintain interconnections, and forces and torques required to maintain velocity constraints.

6.1. Forces and torques distributed on a body. To see how forces and torques of a general character arise, in this section we consider forces and torques as vector measures on a body, and indicate how forces of this type are reduced to point forces and torques. First we give the notion of force and torque as a measure.

6.1 Definition: (Distributed force, distributed torque) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$.

- (i) A **distributed force** on \mathcal{B} at Φ is a vector measure $\sigma: \mathcal{B}(\mathbb{B}) \rightarrow \mathbb{V}$ from the Borel sets of \mathbb{B} such that $\text{supp}(\sigma) \subseteq \mathcal{B}$,
- (ii) A **distributed torque** on \mathcal{B} at Φ is a vector measure $\rho: \mathcal{B}(\mathbb{B}) \rightarrow \mathbb{V}$ such that $\text{supp}(\rho) \subseteq \mathcal{B}$ and

$$\rho(B) \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^\perp_{\mathfrak{g}}, \quad B \in \mathcal{B}(\mathbb{B}). \quad \bullet$$

The condition on a distributed torque means that a torque does no work on the inertialess motions of the system (we will discuss work in Section 6.4). Note that, with distributed forces and torques, we are intermingling body and spatial points of view, since points where force are applied are given in body coordinates (reflected by the fact that a force distribution has as its domain the Borel subsets of \mathbb{B}), but takes values in the spatial vector space \mathbb{V}).

Let us give some examples of distributed forces and torques before we consider further properties and related constructions.

6.2 Examples: (Distributed forces and torques)

1. Let $X_0 \in \mathcal{B}$ and let $f \in \mathbb{V}$. By δ_{X_0} denote the Dirac scalar measure on $\mathcal{B}(\mathbb{B})$. The **point force** f at X_0 is the vector measure $\delta_{X_0} \cdot f$; thus

$$\delta_{x_0} \cdot f(B) = \begin{cases} f, & X_0 \in B, \\ 0, & \text{otherwise,} \end{cases}$$

for $B \in \mathcal{B}(\mathbb{B})$.

2. In like manner, a **point torque** at X_0 is of the form $\delta_{X_0} \cdot \tau$ for some $\tau \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^\perp_{\mathfrak{g}}$.
3. Let us indicate how gravitational forces are distributed forces in our setting. We let $a_g \in \mathbb{V}$ be a nonzero vector in the direction of gravity whose length is the gravitational acceleration in appropriate units. The **gravitational force** on a body (\mathcal{B}, μ) is then the vector measure $F_g = \mu \cdot a_g$. Thus

$$F_g(B) = \mu(B)a_g, \quad B \in \mathcal{B}(\mathbb{B}).$$

4. Let $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ be a C^2 -rigid motion and let (\mathcal{B}, μ) be a rigid body. For $t_0 \in \mathbb{T}'$, we define a force distribution $f_\phi(t_0)$ and a torque distribution $\tau_\phi(t_0)$ by

$$f_\phi(t_0)(B) = \left. \frac{d}{dt} \right|_{t=t_0} \int_B \frac{d}{dt} \phi(t)(X) d\mu(X)$$

and

$$\tau_\phi(t_0)(B) = \left. \frac{d}{dt} \right|_{t=t_0} \int_B (\phi(t)(X) - \phi(t)(X_c)) \times \left(\frac{d}{dt}(\phi(t)(X) - \phi(t)(X_c)) \right) d\mu(X)$$

for $B \in \mathcal{B}(\mathbb{B})$. We call these the *inertial force distribution* and the *inertial torque distribution*. The relevance of these will be apparent in our statement of the Lagrange–d’Alembert Principle in Proposition 7.6. •

With the notion of a point force at hand, we can define some additional constructions with distributed forces.

6.3 Definition: (Resultant force, resultant torque) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . For a distributed force σ and a distributed torque ρ at Φ , and for $x_0 \in \mathbb{S}$:

- (i) the *resultant force* for σ is $f_\sigma = \sigma(\mathbb{B}) \in \mathbb{V}$;
- (ii) the *resultant torque* for σ about x_0 is $\tau_{\sigma, x_0} \in \mathbb{V}$ defined by

$$\tau_{\sigma, x_0} = \int_{\mathbb{B}} (\Phi(X) - x_0) \times d\sigma(X);$$

- (iii) the *resultant torque* for ρ is $\tau_\rho = \rho(\mathbb{B}) \in \mathbb{V}$. •

6.4 Remark: (Computation of resultant torque) Let us clarify the manner in which the resultant torque of a distributed force is computed. Thus we consider a vector measure $\nu: \mathcal{B}(\mathbb{B}) \rightarrow \mathbb{V}$. For $v \in \mathbb{V}$ let

$$p_v: \mathbb{V} \rightarrow \mathbb{R}$$

$$v' \mapsto \begin{cases} \frac{\mathfrak{g}(v', v)}{\mathfrak{g}(v, v)}, & v \neq 0, \\ 0, & v = 0. \end{cases}$$

For $v \in \mathbb{V}$ we define a signed measure ν_v on \mathbb{B} by $\nu_v = p_v \circ \nu$. With this notation, given an oriented orthonormal basis (e_1, e_2, e_3) for \mathbb{V} , we then define

$$\int_{\mathbb{B}} (\Phi(X) - x_0) \times d\nu(X) \triangleq \sum_{j=1}^3 \int_{\mathbb{B}} (\Phi(X) - x_0) \times e_j d\nu_{e_j}(X),$$

understanding, as in Remark 3.3, how we integrate functions taking values in finite-dimensional vector spaces. It is a simple matter to show that this definition is independent of the choice of basis. •

With these constructions, we have the following notion.

6.5 Definition: (Centre of force) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . A *centre of force* for a distributed force σ at $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ is a point $x_\sigma \in \mathbb{S}$ satisfying $\tau_{\sigma, x_\sigma} = 0$. •

The following theorem details the character of centres of force.

6.6 Theorem: ((Non)existence and (non)uniqueness of centre of force) *Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ and let σ be a distributed force on \mathcal{B} at Φ . The following statements hold:*

(i) *if $f_\sigma \neq 0$, then a centre of force exists and, if x_σ is a centre of force, then the set of all centres of force is*

$$\{x_\sigma + af_\sigma \mid a \in \mathbb{R}\};$$

(ii) *if $f_\sigma = 0$, then there exists $\tau_\sigma \in \mathbf{V}$ such that $\tau_{\sigma, x_0} = \tau_\sigma$ for every $x_0 \in \mathbf{S}$, and we have the following cases:*

(a) *if $\tau_\sigma \neq 0$ then there is no centre of force;*

(b) *if $\tau_\sigma = 0$ then the set of all centres of force is \mathbf{S} .*

Proof: Let $x_0 \in \mathbf{S}$, let $\bar{x} \in \mathbf{S}$, and note that \bar{x} is a centre of force if and only if

$$\begin{aligned} & \int_{\mathcal{B}} (\Phi(X) - \bar{x}) \times d\sigma(X) = 0 \\ \iff & (\bar{x} - x_0) \times \int_{\mathcal{B}} d\sigma(X) = \int_{\mathcal{B}} (\Phi(X) - x_0) \times d\sigma(X). \end{aligned}$$

We note that this is a linear equation $A_\sigma(\bar{x} - x_0) = b_{\sigma, x_0}$ for $\bar{x} - x_0 \in \mathbf{V}$, where

$$A_\sigma(v) = -f_\sigma \times v, \quad b_{\sigma, x_0} = \int_{\mathcal{B}} (\Phi(X) - x_0) \times d\sigma(X).$$

We examine this equation in various cases.

First we suppose that $f_\sigma \neq 0$. We claim that $\ker(A_\sigma) = \text{image}(A_\sigma)^{\perp \mathbf{g}}$. Indeed,

$$\begin{aligned} \ker(A_\sigma) &= \{v \mid A_\sigma(v) = 0\} \\ &= \{v \mid \mathbf{g}(A_\sigma(v), v') = 0, v' \in \mathbf{V}\} \\ &= \{v \mid \mathbf{g}(f_\sigma \times v, v') = 0, v' \in \mathbf{V}\} \\ &= \{v \mid \mathbf{g}(f_\sigma \times v', v) = 0, v' \in \mathbf{V}\} \\ &= \text{image}(A_\sigma)^{\perp \mathbf{g}}, \end{aligned}$$

using (1.1). Since $\ker(A_\sigma) = \text{span}(f_\sigma)$, the linear equation for $\bar{x} - x_0$ has solutions if and only if b_{σ, x_0} is \mathbf{g} -orthogonal to f_σ . To this end, we let (e_1, e_2, e_3) be an oriented orthonormal basis for \mathbf{V} and compute

$$\begin{aligned} \mathbf{g}(f_\sigma, b_{\sigma, x_0}) &= \mathbf{g}\left(\int_{\mathcal{B}} d\sigma(X), \int_{\mathcal{B}} (\Phi(X') - x_0) \times d\sigma(X')\right) \\ &= \sum_{j,k=1}^3 \int_{\mathcal{B}} \left(\int_{\mathcal{B}} \mathbf{g}(e_j, (\Phi(X') - x_0) \times e_k) d\sigma_{e_k}(X')\right) d\sigma_{e_j}(X) \\ &= - \sum_{j,k=1}^3 \int_{\mathcal{B}} \left(\int_{\mathcal{B}} \mathbf{g}(e_k, (\Phi(X') - x_0) \times e_j) d\sigma_{e_k}(X')\right) d\sigma_{e_j}(X) \\ &= - \sum_{j,k=1}^3 \int_{\mathcal{B}} \left(\int_{\mathcal{B}} \mathbf{g}(e_k, (\Phi(X') - x_0) \times e_j) d\sigma_{e_j}(X')\right) d\sigma_{e_k}(X) \\ &= - \mathbf{g}(f_\sigma, b_{\sigma, x_0}), \end{aligned}$$

using (1.1) for the third line and Fubini's Theorem for the fourth line. This shows that b_{σ, x_0} is indeed g -orthogonal to f_σ , showing that the linear equation for $\bar{x} - x_0$ has solutions when $f_\sigma \neq 0$. Moreover, any two such solutions will differ by an element of $\ker(A_\sigma)$, and this proves (i).

Now suppose that $f_\sigma = 0$ and let $x_1, x_2 \in \mathcal{S}$. We compute

$$\begin{aligned}\tau_{\sigma, x_2} &= \int_{\mathcal{B}} (\Phi(X) - x_2) \times d\sigma(X) \\ &= \int_{\mathcal{B}} (\Phi(X) - x_1) \times d\sigma(X) + (x_1 - x_2) \times \int_{\mathcal{B}} d\sigma(X) = \tau_{\sigma, x_1},\end{aligned}$$

showing that the resultant torque τ_{σ, x_0} is indeed independent of x_0 , which gives the first assertion in part (ii). Let us denote the resultant torque, then, by τ_σ . First suppose that $\tau_\sigma \neq 0$. Then, in our linear equation above for $\bar{x} - x_0$, $A_\sigma = 0$ and $b_{\sigma, x_0} \neq 0$. Thus the linear equation has no solutions. When $\tau_\sigma = 0$, the linear equation simply reads “ $0 = 0$,” and so the set of solutions is all of \mathcal{V} . \blacksquare

We are not aware of this sort of result having been stated or proved in this generality, so let us make a few comments on it.

6.7 Remarks: (Centre of force)

1. The fact that, when $f_\sigma \neq 0$, the centre of force is not uniquely defined is a reflection of the fact that a force can be moved along its line of application and the resultant torque, which is orthogonal to this line, remains unchanged.
2. When $f_\sigma = 0$ but $\tau_\sigma \neq 0$, then the force is a “pure torque.” The fact that, in this case, a centre of force does not exist is a reflection of the oft-cited principle that torque is not applied about any particular point.
3. When both f_σ and τ_σ vanish, then the force is an equilibrium force. That the centre of force is arbitrary is a reflection of the oft-cited principle that, at equilibrium, one can take moments of forces about any point one wishes. \bullet

Let us illustrate the preceding notions for the force distributions we introduced in Example 6.2.

6.8 Examples: (Resultant force, resultant torque, centre of force)

1. For the point force f at X_0 , the resultant force is obviously f and the set of centres of force is

$$\{\Phi(X_0) + \alpha f \mid \alpha \in \mathbb{R}\}.$$

2. It is an easy matter to show that the centre of mass is a centre of force for the gravitational force distribution of Example 6.2–3. The set of centres of force is

$$\{\Phi(X_c) + \alpha a_g \mid \alpha \in \mathbb{R}\}.$$

The resultant force is, of course, ma_g .

3. In Example 6.2–4 we introduce the inertial force for a body (\mathcal{B}, μ) along a motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ at $t_0 \in \mathbb{T}'$. The set of centres of force for this force distribution is easily seen to be

$$\{\Phi(X_c) + \alpha \dot{v}_\phi(t_0) \mid \alpha \in \mathbb{R}\}.$$

The resultant force is $m\dot{v}_\phi(t_0)$, as we shall see in Lemma 7.1. The resultant torque is computed to be

$$\left. \frac{d}{dt} \right|_{t=t_0} R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)),$$

as we shall see in Lemma 7.1. •

We can now reduce a distributed force to a specific, equivalent point force and a torque.

6.9 Definition: (Central force and torque) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$, and let σ be a distributed force and ρ be a distributed torque on \mathcal{B} at Φ .

- (i) The *central force* for σ is the point force $F_\sigma = \delta_{X_c} \cdot f_\sigma$.
- (ii) The *central torque* for σ is $\tau_\sigma = \tau_{\sigma, \Phi(X_c)}$.
- (iii) The *central torque* for ρ is the resultant torque τ_ρ . •

The fact that the central force is a force distribution, whereas the central torques are merely vectors in \mathbb{V} , reflects the fact that forces require points of application, whereas torques do not.

We should verify that central torques do no work on inertialess motions.

6.10 Lemma: (Central torques and inertialess motions) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$. If σ is a force distribution and ρ is a torque distribution on \mathcal{B} at Φ , then

$$\tau_\sigma, \tau_\rho \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^{\perp_{\mathfrak{g}}}.$$

Proof: First we show that

$$\tau_\sigma \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^{\perp_{\mathfrak{g}}}.$$

We consider three cases. First, if the inertia tensor is nondegenerate, the assertion follows trivially since $(\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^{\perp_{\mathfrak{g}}} = \mathbb{V}$. If the inertia tensor is degenerate but nonzero, then all points on the body lie on a line through X_c in the direction of $\ker(\mathbb{I}_c)$ (combining Example 4.4–2 and the proof of Lemma 4.8). Let (e_1, e_2, e_3) be an oriented orthogonal basis for \mathbb{V} . If $X \in \mathcal{B}$, then

$$X - X_c \in \ker(\mathbb{I}_c) \implies R_\Phi(X - X_c) \in \ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T).$$

Thus, for any $X \in \mathcal{B}$ and $v \in \mathbb{V}$, we have

$$(\Phi(X) - \Phi(X_c)) \times v \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^{\perp_{\mathfrak{g}}}.$$

Then we have, for every $v \in \ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T)$,

$$\mathfrak{g}(v, \tau_\sigma) = \sum_{j=1}^3 \int_{\mathcal{B}} \mathfrak{g}(v, (\Phi(X) - \Phi(X_c)) \times e_j) d\sigma_{e_j}(X) = 0,$$

giving the assertion in this case. Finally, if the inertia tensor is zero, then the body is a particle, and so $X - X_c = 0$ for every $X \in \mathcal{B}$. Since $(\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^{\perp_{\mathfrak{g}}} = \{0\}$, our assertion follows immediately in this case.

That $\tau_\rho \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^{\perp_{\mathfrak{g}}}$ follows by definition. ■

As we shall see in Theorem 7.4—and as is well-known—a distributed force can be replaced with its central force and torque and the dynamics will be unchanged. For this reason, we shall often use a point force and torque in place of a distributed force. When the effects of forces arising from multiple physical effects are to be accumulated, this can simply be accounted for by adding the force distributions by the following reasoning: (1) the set of distributed forces is a vector space under pointwise addition and scalar multiplication (with points being elements of the Borel σ -algebra); (2) the maps assigning a resultant force and torque to a distributed force are \mathbb{R} -linear.

6.2. Primary torque-force. With the above construction of forces on physical grounds, and the reduction of these to central forces and torques, we next give the elementary mathematical definition of forces as we shall use them to incorporate into our simple equations of motion for a rigid body.

6.11 Definition: (Primary torque-force (single rigid body)) Let $\mathcal{S} = (\mathcal{S}, \mathcal{V}, g, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, G, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , and let $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$. A *primary torque-force* on \mathcal{B} at Φ is a pair $(\tau, f) \in \mathcal{V} \oplus \mathcal{V}$ such that

$$\tau \in (\ker(R_\Phi \circ \mathbb{I}_c \circ R_\Phi^T))^\perp. \quad \bullet$$

Mathematically, of course, these definitions merely say that central force and torque are elements of \mathcal{V} . The condition on a primary torque indicates that a torque should do no work on the inertialess motions of the system. In accord with our discussions regarding distributed forces, we think always of a primary force as being a point force at the centre of mass, or more precisely at the image of the centre of mass under the configuration Φ . We think of the direction of a primary torque τ as being the axis about which the torque is applied and the magnitude of the torque is the length of τ . In Figure 3 we depict how

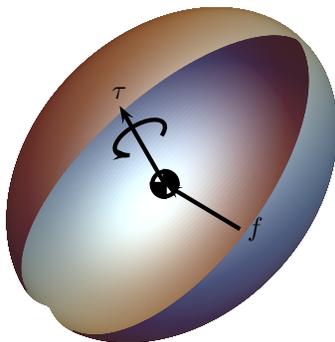


Figure 3. Central torque-force on a rigid body in a configuration

one should physically think of a primary force and torque.

It will be convenient to extend the previous single body construction to systems of rigid bodies.

6.12 Definition: (Primary torque-force (multiple rigid bodies)) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathcal{Q} . Let $\Phi \in \mathcal{Q}$. A *primary torque-force* on \mathcal{B} at Φ is a pair $(\tau, \mathbf{f}) \in \mathbb{V}^m \oplus \mathbb{V}^m$ such that (τ_a, \mathbf{f}_a) is a primary torque force on \mathcal{B}_a at Φ_a . •

Note that a primary torque-force can arise from an amalgamation of force and torque distributions because of the additivity of distributed forces and torques. Thus a primary torque-force is to be thought of as a simple mathematical representation of what may well be complicated physical data.

6.3. Torque-force fields. In practice, external (rather than internal, which we shall consider shortly) forces and torques are not just prescribed at a point and in a given configuration, but are given as functions of, possibly, time, position, and velocity. Other dependencies are possible and can be worked into the constructions as needed.

We first consider the case of force and torque distributions on single rigid bodies. To facilitate this, we denote by $\mathcal{M}(\mathcal{B}; \mathbb{V})$ the set of vector measures from the Borel sets $\mathcal{B}(\mathcal{B})$ into \mathbb{V} .

6.13 Definition: (Force distribution field, torque distribution field) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}, \mathbb{U}, \mathbf{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , let \mathbb{T} be a time axis, and let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval.

(i) A *force distribution field* is a mapping

$$\sigma: \mathbb{T}' \times \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \mathcal{M}(\mathcal{B}; \mathbb{V})$$

and

(ii) a *torque distribution field* is a mapping

$$\rho: \mathbb{T}' \times \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \mathcal{M}(\mathcal{B}; \mathbb{V})$$

for which $\rho(t, (\Phi, (A, v)))$ is a torque distribution for \mathcal{B} at Φ for each $(t, (\Phi, (A, v))) \in \mathbb{T}' \times \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$. •

Of course, given a force and torque distribution field and $(t, (\Phi, (A, v))) \in \mathbb{T}' \times \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$, there are central force and torque distributions associated to the force and torque distributions $\sigma(t, (\Phi, (A, v)))$ and $\rho(t, (\Phi, (A, v)))$. Thus, as usual, the effects of force distribution and torque distribution fields can be reduced to the consideration of primary force and torque fields, a notion that we define next.

6.14 Definition: (Torque-force field) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, configuration manifold \mathcal{Q} , time axis \mathbb{T} , and time interval $\mathbb{T}' \subseteq \mathbb{T}$. A *torque-force field* is a mapping

$$\begin{aligned} \tau \oplus \mathbf{f}: \mathbb{T}' \times \mathbb{T}\mathcal{Q} &\rightarrow \mathbb{V}^m \oplus \mathbb{V}^m \\ (t, (\Phi, (A, v))) &\mapsto (\tau(t, (\Phi, (A, v))), \mathbf{f}(t, (\Phi, (A, v)))) \end{aligned}$$

such that

$$\tau \oplus \mathbf{f}(t, \alpha_{\mathcal{B}, \Psi}(\Phi, (A, v))) = \tau \oplus \mathbf{f}(t, (\Phi, (A, v)))$$

for all $\Psi \in l(\mathcal{B})$ and such that, for each $t \in \mathbb{T}'$, $(\Phi, (\mathbf{A}, \mathbf{v})) \in \text{TQ}$, and $a \in \{1, \dots, m\}$,

$$(\tau_a(t, (\mathbf{A}, \mathbf{v})), f_a(t, (\mathbf{A}, \mathbf{v})))$$

is a primary torque-force on \mathcal{B}_a at Φ_a . •

The preceding definition of torque-force fields is very general, and will not have any useful properties, e.g., one may not have a motion associated to the physical laws governing rigid body motion, i.e., the Newton–Euler equations that we will consider in Section 7. Since such motions arise, of course, as solutions to differential equations, the regularity with which torque-force fields depend on time and position/velocity can be as general as required for existence and uniqueness of solutions to differential equations. This is considered in a unified and general way, for many sorts of regularity, in the book of [Jafarpour and Lewis \[2014\]](#). Here we will not burden our presentation with such generality, and shall merely say that a torque-force field is *smooth* if it is a smooth function on $\mathbb{T}' \times \text{TQ}$ in the usual sense. If we have force distribution and torque distribution fields defined on the bodies of our system, we shall say they are *weakly smooth* if the resulting torque-force field, defined using the associated central force and torque distribution fields, is smooth.

Let us now show that, in the presence of internal symmetry, torque-force fields descend to the reduced configuration manifold. First we denote

$$\text{T-F} = \{(\Phi, \tau \oplus \mathbf{f}) \in \mathcal{Q} \times (\mathbf{V}^m \oplus \mathbf{V}^m) \mid \tau_a \in (\ker(R_{\Phi_a} \circ \mathbb{I}_{c,a} \circ R_{\Phi}^T))^\perp, a \in \{1, \dots, m\}\},$$

which we think of as being the vector bundle where torque-force fields take their values. This vector bundle has an action of $l(\mathcal{B})$, this being given by

$$((\Phi, \tau \oplus \mathbf{f}), \Psi) \mapsto (\alpha_{\mathcal{B}}(\Phi, \Psi), \tau \oplus \mathbf{f}),$$

cf. equation (4.2). Note that, by Lemma 4.6, the fibre of T-F over $\Phi \in \mathcal{Q}$ depends only on the $l(\mathcal{B})$ -orbit $\pi_0(\Phi)$. Thus T-F/ $l(\mathcal{B})$ is a well-defined vector bundle over \mathcal{Q}_0 and the projection to the quotient has the form

$$(\Phi, \tau \oplus \mathbf{f}) \mapsto (\pi_0(\Phi), \tau \oplus \mathbf{f}), \tag{6.1}$$

with appropriate restrictions on τ .

With the preceding as backdrop, we have the following lemma.

6.15 Lemma: (Torque-force fields descend to the reduced configuration manifold) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. If $\tau \oplus \mathbf{f}$ is a torque-force field, then the following statements hold:*

(i) *there exist a mapping*

$$\tau_0 \oplus \mathbf{f}_0: \mathbb{T}' \times \text{TQ}_0 \rightarrow \text{T-F}_{\text{spatial}}/l(\mathcal{B})$$

satisfying

$$\tau_0 \oplus \mathbf{f}_0(t, T\pi_0(\Phi, (\mathbf{A}, \mathbf{v}))) = (\pi_0(\Phi), \tau(t, (\Phi, (\mathbf{A}, \mathbf{v}))), \mathbf{f}(t, (\Phi, (\mathbf{A}, \mathbf{v}))))); \tag{6.2}$$

(ii) if $\boldsymbol{\tau} \oplus \boldsymbol{f}$ is smooth, then $\boldsymbol{\tau}_0 \oplus \boldsymbol{f}_0$ is smooth.

Proof: (i) The existence and well-definedness of $\boldsymbol{\tau}_0 \oplus \boldsymbol{f}_0$ follows directly from the invariance of $\boldsymbol{\tau} \oplus \boldsymbol{f}$ under the action of $l(\mathcal{B})$. That $\boldsymbol{\tau}_0 \oplus \boldsymbol{f}_0$ satisfies (6.2) is a consequence of the fact that points in T-F/ $l(\mathcal{B})$ admit the representation (6.1).

(ii) This follows since π_0 and $T\pi_0$ are smooth surjective submersions by Proposition 4.9. ■

The final notion we define is that of a torque-force field along a motion of a system.

6.16 Definition: (Torque-force along a motion) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. For a differentiable motion $\phi: \mathbb{T}' \rightarrow \mathcal{Q}$, a **torque-force field along ϕ** is a mapping $\boldsymbol{\tau} \oplus \boldsymbol{f}: \mathbb{T}' \rightarrow \mathbb{V}^m \oplus \mathbb{V}^m$ such that $(\boldsymbol{\tau}(t), \boldsymbol{f}(t))$ is a primary torque-force on \mathcal{B} at $\phi(t)$. •

6.4. Work and power. An important physical notion that can be used to assign attributes to forces and torques is that of work, and its infinitesimal variant, power. In this section we indicate how these concepts are defined in our setting.

6.17 Definition: (Work, power) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. For a differentiable motion $\phi: \mathbb{T}' \rightarrow \mathcal{Q}$ and for a continuous torque-force field $\boldsymbol{\tau} \oplus \boldsymbol{f}: \mathbb{T}' \rightarrow \mathbb{V}^m \oplus \mathbb{V}^m$ along ϕ :

(i) the pair $(\boldsymbol{\tau} \oplus \boldsymbol{f}, \phi)$ is **integrable** if the integrals

$$\int_{\mathbb{T}'} g(\tau_a(t), \omega_{\phi_a}(t)) dt, \quad \int_{\mathbb{T}'} g(f_a(t), v_{\phi_a}(t)) dt, \quad a \in \{1, \dots, m\},$$

exist and are finite;

(ii) if $(\boldsymbol{\tau} \oplus \boldsymbol{f}, \phi)$ is integrable, the **work** done by $\boldsymbol{\tau} \oplus \boldsymbol{f}$ along ϕ is

$$W(\boldsymbol{\tau} \oplus \boldsymbol{f}, \phi) = \sum_{a=1}^m \int_{\mathbb{T}'} g(\tau_a(t), \omega_{\phi_a}(t)) dt + \sum_{a=1}^m \int_{\mathbb{T}'} g(f_a(t), v_{\phi_a}(t)) dt;$$

(iii) the **power** of $\boldsymbol{\tau} \oplus \boldsymbol{f}$ along ϕ is

$$\begin{aligned} P(\boldsymbol{\tau} \oplus \boldsymbol{f}, \phi): \mathbb{T}' &\rightarrow \mathbb{R} \\ t &\mapsto \sum_{a=1}^m g(\tau_a(t), \omega_{\phi_a}(t)) + \sum_{a=1}^m g(f_a(t), v_{\phi_a}(t)). \end{aligned} \quad \bullet$$

Work and power can be dropped to the reduced configuration space. Note that we have the following natural bilinear pairing on the fibres of vector bundles:

$$\begin{aligned} \Sigma_g: \text{T-F}/l(\mathcal{B}) \times \mathbb{V}_{\text{spatial}}^{\text{red}}|_{\mathcal{Q}_0} &\rightarrow \mathbb{R} \\ ((\pi_0(\Phi), \boldsymbol{\tau} \oplus \boldsymbol{f}), [(\Phi, \boldsymbol{\omega} \oplus \boldsymbol{v}) + \mathbb{Z}_{\text{spatial}}]) &\mapsto \sum_{a=1}^m (g(\tau_a, \omega_a) + g(f_a, v_a)). \end{aligned}$$

That this pairing is well-defined follows immediately from the definitions of T-F and $\mathbb{V}_{\text{spatial}}^{\text{red}}$ and because the fibres of T-F/I(\mathcal{B}) are “the same” as the fibres of T-F, cf. (6.1).

6.18 Lemma: (Reduction of work and power) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathbb{Q} . Let \mathbb{T} be a time axis, let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval, let $\phi: \mathbb{T}' \rightarrow \mathbb{Q}$ be a differentiable curve, and let $\tau \oplus \mathbf{f}: \mathbb{T}' \rightarrow \mathbb{V}^m \oplus \mathbb{V}^m$ be a continuous torque-force field along ϕ . Then*

$$P(\tau \oplus \mathbf{f}, \phi)(t) = \Sigma_{\mathbf{g}}((\phi_0(t), \tau \oplus \mathbf{f}(t)), [(\phi(t), \boldsymbol{\omega}_{\phi} \oplus \mathbf{v}_{\phi}(t)) + \mathbb{Z}_{\text{spatial}}]), \quad t \in \mathbb{T}'.$$

and, if $(\tau \oplus \mathbf{f}, \phi)$ is integrable,

$$W(\tau \oplus \mathbf{f}, \phi) = \int_{\mathbb{T}'} \Sigma_{\mathbf{g}}((\phi_0(t), \tau \oplus \mathbf{f}(t)), [(\phi(t), \boldsymbol{\omega}_{\phi} \oplus \mathbf{v}_{\phi}(t)) + \mathbb{Z}_{\text{spatial}}]) dt.$$

Proof: Clearly it suffices to prove the assertion for power, but this follows directly from the definition of $\Sigma_{\mathbf{g}}$. ■

6.5. Virtual displacements, virtual power, and virtual work. The Principle of Virtual Work, and the closely related Lagrange–d’Alembert Principle, are devices for characterising torque-forces and for understanding the rôles of torque-force in variational principles. Behind these principles are the notions of virtual displacement and virtual work. The presentation of these ideas in the applied literature is typically very opaque and, in the mathematical literature, they are typically not covered at all, or at best as a mathematical triviality.⁴ Therefore, here we provide a clear characterisation of the tools used in the Principle of Virtual Work, including the physical ideas behind these.

We consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathbb{Q} . We let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval, and let $\phi: \mathbb{T}' \rightarrow \mathbb{Q}$ be a differentiable motion for the system. We then consider a variation of this motion, i.e., a differentiable map

$$\sigma: [-r, r] \times \mathbb{T}' \rightarrow \mathbb{Q}_{\text{free}}$$

such that $\sigma(0, t) = \phi(t)$. Let us denote $\sigma_s = \sigma(s, t)$ and define $\sigma'(s, t) = \sigma'_s(t)$. Note that σ_s is a curve on \mathbb{Q} for all $s \in [-r, r]$. Denote

$$\delta\sigma(t) = \left. \frac{d}{ds} \right|_{s=0} \sigma(s, t).$$

By choosing a spatial origin x_0 we can define, for $a \in \{1, \dots, m\}$,

$$R_{\sigma_a}(s, t) = R_{\sigma_a(s, t)}, \quad r_{\sigma_a}(s, t) = r_{\sigma_a(s, t)},$$

and so

$$\delta_{\sigma_a} R_{\sigma_a}(t) = \left. \frac{d}{ds} \right|_{s=0} R_{\sigma_a}(s, t), \quad \delta_{\sigma_a} r_{\sigma_a}(t) = \left. \frac{d}{ds} \right|_{s=0} r_{\sigma_a}(s, t).$$

We then denote

$$\boldsymbol{\omega}_{\sigma_a}(t) = \delta_{\sigma_a} \widetilde{R_{\sigma_a}(t)} \circ R_{\sigma_a}^T(0, t), \quad v_{\sigma_a}(t) = \delta_{\sigma_a} r_{\sigma_a}(t).$$

With these constructions at hand, let us make the basic definitions we will use.

⁴A reader interested in a detailed mathematical presentation of these principles should look at their treatment by Spivak [Spivak 2010].

6.19 Definition: (Virtual displacement, virtual power, virtual work) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Suppose that we have velocity constraints defining a constraint distribution \mathcal{D} . We let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval, and let $\phi: \mathbb{T}' \rightarrow \mathcal{Q}$ be a differentiable motion for the system for which $\phi'(t) \in \mathcal{D}_{\phi(t)}$ for each $t \in \mathbb{T}'$. Let $\tau \oplus \mathbf{f}$ be a torque-force field along ϕ . With the above constructions, we have the following definitions.

- (i) A **variation** of ϕ is a differentiable mapping $\sigma: [-r, r] \times \mathbb{T}' \rightarrow \mathcal{Q}_{\text{free}}$ for some $r \in \mathbb{R}_{>0}$ such that $\sigma(0, t) = \phi(t)$ for all $t \in \mathbb{T}'$.
- (ii) A **Q-variation** is a variation for which $\sigma(s, t) \in \mathcal{Q}$ for all $(s, t) \in [-r, r] \times \mathbb{T}'$.
- (iii) A **D-variation** is a Q-variation for which $\sigma'(s, t) \in \mathcal{D}_{\sigma(s, t)}$ for all $(s, t) \in [-r, r] \times \mathbb{T}'$.
- (iv) The **virtual displacement** associated with a variation σ is the map $\delta\sigma: \mathbb{T}' \rightarrow \mathcal{T}\mathcal{Q}_{\text{free}}$.
- (v) An **infinitesimal Q-variation** is a variation for which $\delta\sigma(t) \in \mathcal{T}_{\phi(t)}\mathcal{Q}$ for all $t \in \mathbb{T}'$.
- (vi) An **infinitesimal D-variation** is a variation for which $\delta\sigma(t) \in \mathcal{D}_{\phi(t)}$ for all $t \in \mathbb{T}'$.
- (vii) The **virtual power** of $\tau \oplus \mathbf{f}$ associated to the virtual displacement $\delta\sigma$ is

$$t \mapsto \sum_{a=1}^m (g(\tau_a(t), \omega_{\sigma_a}(t)) + g(f_a(t), v_{\sigma_a}(t))).$$

- (viii) If $(\tau \oplus \mathbf{f}, \phi)$ is integrable, the **virtual work** done by $\tau \oplus \mathbf{f}$ along ϕ associated to the virtual displacement $\delta\sigma$ is

$$\sum_{a=1}^m \int_{\mathbb{T}'} g(\tau_a(t), \omega_{\sigma_a}(t)) dt + \sum_{a=1}^m \int_{\mathbb{T}'} g(f_a(t), v_{\sigma_a}(t)) dt. \quad \bullet$$

We note that a virtual displacement is a vector field with values in $\mathcal{T}\mathcal{Q}_{\text{free}}$ along the curve ϕ . An important question is just which vector fields along ϕ arise as virtual displacements. The following result provides some answers to this question.

6.20 Proposition: (The character of virtual displacements) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Suppose that we have velocity constraints defining a constraint distribution \mathcal{D} . We let $\phi: [t_0, t_1] \rightarrow \mathcal{Q}$ be a differentiable motion for the system for which $\phi'(t) \in \mathcal{D}_{\phi(t)}$ for each $t \in [t_0, t_1]$. The following statements hold:

- (i) if $\mathbf{V}: [t_0, t_1] \rightarrow \mathcal{T}\mathcal{Q}_{\text{free}}$ is a differentiable mapping such that $\mathbf{V}(t) \in \mathcal{T}_{\phi(t)}\mathcal{Q}_{\text{free}}$, then there exists a variation σ of ϕ such that $\delta\sigma = \mathbf{V}$;
- (ii) if $\mathbf{V}: [t_0, t_1] \rightarrow \mathcal{T}\mathcal{Q}_{\text{free}}$ is a differentiable mapping such that $\mathbf{V}(t) \in \mathcal{T}_{\phi(t)}\mathcal{Q}$, then there exists a Q-variation σ of ϕ such that $\delta\sigma = \mathbf{V}$;
- (iii) if σ is a Q-variation, then it is an infinitesimal Q-variation;
- (iv) there exists a constraint distribution \mathcal{D} and a C^1 -mapping $\mathbf{V}: [t_0, t_1] \rightarrow \mathcal{T}\mathcal{Q}_{\text{free}}$ such that $\mathbf{V}(t) \in \mathcal{D}_{\phi(t)}$ and such that there is no \mathcal{D} -variation σ of class C^2 for which $\delta\sigma = \mathbf{V}$;
- (v) there exist a constraint distribution \mathcal{D} and a \mathcal{D} -variation that is not an infinitesimal \mathcal{D} -variation.

Proof: (i) We shall prove this in a more general setting in Lemma 8.3 below. Note that, while the vector field in Lemma 8.3 vanishes at the endpoints of the interval, the proof does not rely on this fact as it only ensures that the resulting variation is a fixed endpoint variation, which we are not requiring here.

(ii) This also follows using the arguments from the proof of Lemma 8.3.

(iii) This is obvious since the curve $s \mapsto \sigma(s, t)$ is in \mathbf{Q} if σ is a \mathbf{Q} -variation.

(iv) We will build an example. We will work with a single body and take $\mathbf{U} = \mathbf{V} = \mathbb{R}^3$ with $\mathbf{S} = \mathbf{B} = \mathbb{R}^3$ affine spaces in the canonical way. We use coordinates (x, y, z) for \mathbb{R}^3 . We thus have $\mathbf{Q}_{\text{free}} = \text{SO}(3) \times \mathbb{R}^3$ and we define a distribution on \mathbf{Q}_{free} by defining a distribution on \mathbb{R}^3 , then asking that it have no component in the $\text{SO}(3)$ -direction. The distribution \mathbf{D} that we define on \mathbb{R}^3 is that generated by the vector fields

$$X = \frac{\partial}{\partial x}, \quad Y = \frac{\partial}{\partial y} + x \frac{\partial}{\partial z}.$$

We take a curve

$$\begin{aligned} \phi: [0, 1] &\rightarrow \mathbb{R}^3 \\ t &\mapsto (0, t, 0) \end{aligned}$$

that satisfies $\phi'(t) \in \mathbf{D}_{\phi(t)}$. We define a vector field

$$\begin{aligned} \mathbf{V}: [0, 1] &\rightarrow \mathbb{T}\mathbb{R}^3 \\ t &\mapsto \frac{\partial}{\partial x} \end{aligned}$$

along ϕ , noting that $\mathbf{V}(t) \in \mathbf{D}_{\phi(t)}$ for all $t \in [0, 1]$. Now consider a variation $\sigma: [-r, r] \times [0, 1] \rightarrow \mathbb{R}^3$ of ϕ and write

$$\sigma(s, t) = (\sigma_x(s, t), \sigma_y(s, t), \sigma_z(s, t)).$$

One readily sees that σ is a \mathbf{D} -variation if and only if

$$\dot{\sigma}_z(s, t) - \sigma_x(s, t)\dot{\sigma}_y(s, t) = 0, \quad (s, t) \in [-r, r] \times [0, 1],$$

“dot” representing time-derivative. Differentiating this expression with respect to s at $s = 0$, and using equality of mixed partials, we have

$$\frac{d}{dt}\delta\sigma_z(t) - \delta\sigma_x(t)\frac{d}{dt}\sigma_y(0, t) - \sigma_x(0, t)\frac{d}{dt}\delta\sigma_y(t) = 0.$$

Since

$$\frac{d}{dt}V_z(t) - V_x(t)\frac{d}{dt}\sigma_y(0, t) - \sigma_x(0, t)\frac{d}{dt}V_y(t) = 1,$$

we conclude that there is no \mathbf{D} -variation σ of class \mathbf{C}^2 for which $\mathbf{V} = \delta\sigma$.

Now let us see how we can conclude the same thing for variations, not just in \mathbb{R}^3 , but in $\text{SO}(3) \times \mathbb{R}^3$. This follows since, by the reasoning that gives the conclusion of part (ii) of the proposition, the following two statements are equivalent:

1. there exists a \mathbf{D} -variation, in $\text{SO}(3) \times \mathbb{R}^3$, of ϕ of class \mathbf{C}^2 such that $\delta\sigma = \mathbf{V}$;
2. there exists a \mathbf{D} -variation, in $\{\mathbf{id}\} \times \mathbb{R}^3$, of ϕ of class \mathbf{C}^2 such that $\delta\sigma = \mathbf{V}$.

(v) We do as in the previous part of the proof and work with $\mathbb{Q}_{\text{free}} = \text{SO}(3) \times \mathbb{R}^3$ and consider the same distribution \mathbf{D} . We consider the same curve ϕ as in the previous part of the proof. We then consider the \mathbf{D} -variation

$$\begin{aligned} \sigma: [-1, 1] \times [0, 1] &\rightarrow \mathbb{R}^3 \\ (s, t) &\mapsto (s, t, st). \end{aligned}$$

We then compute

$$\delta\sigma(t) = \frac{\partial}{\partial x} + t \frac{\partial}{\partial z} \in \mathbb{T}_{(0,t,0)}\mathbb{R}^3,$$

which we readily determine is not \mathbf{D} -valued. In this case, the extension of the conclusion to $\text{SO}(3) \times \mathbb{R}^3$ is immediate, since one can consider the variation

$$(s, t) \mapsto (\mathbf{id}, (s, t, st)) \in \text{SO}(3) \times \mathbb{R}^3$$

of the curve $t \mapsto (\mathbf{id}, (0, t, 0))$ in $\text{SO}(3) \times \mathbb{R}^3$. ■

As we shall see in Section 6.7, the problems of parts (iv) and (v) of the preceding result have real consequences in mechanics, consequences that have caused significant confusion in the literature dealing with velocity constraints.

6.6. Interconnection torque-forces. Interconnected rigid body systems require interconnection forces and torques between the bodies and with the external surroundings to maintain membership in the configuration manifold. To this end, we make the following definition.

6.21 Definition: (Interconnection torque-force) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathbb{Q} . Let $\Phi \in \mathbb{Q}$. A primary torque-force $(\iota, \mathbf{i}) \in \mathbb{V}^m \oplus \mathbb{V}^m$ is an *interconnection torque-force* at Φ if $P(\iota \oplus \mathbf{i}, \psi)(0) = 0$ for every differentiable motion $\psi: [-r, r] \rightarrow \mathbb{Q}_{\text{free}}$ of the system satisfying $\psi'(0) \in \mathbb{T}_{\Phi}\mathbb{Q}$. We denote

$$\text{IT-F} = \{(\Phi, \tau \oplus \mathbf{f}) \in \mathbb{T}\text{-F} \mid (\tau, \mathbf{f}) \text{ is an interconnection torque-force at } \Phi\},$$

which is the *interconnection torque-force bundle* for the system. •

Let us explore the relationship between our characterisation of interconnection torque-forces and virtual displacements. By the very definition, interconnection torque-forces are those that have zero virtual power for virtual displacements associated to infinitesimal \mathbb{Q} -variations, i.e., virtual displacements that are tangent to \mathbb{Q} . By parts (i) and (ii) of Proposition 6.20, we can work with either infinitesimal \mathbb{Q} -variations or with \mathbb{Q} -variations. By taking integrals along motions, we can also say that we have zero virtual work for these virtual displacements along motions of the system. The significance of using virtual work in place of virtual power is not revealed here, but only when we discuss variational principles in Section 8.

6.7. Constraint torque-forces. We next consider forces and torques required to maintain velocity constraints. As with interconnection forces and torques, the crucial ingredient is the power of the forces and torques along motions of the system.

6.22 Definition: (Constraint forces and torques) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Suppose that we have physical velocity constraints giving rise to a smooth locally constant rank constraint distribution \mathcal{D} . Let $\Phi \in \mathcal{Q}$. A primary torque-force $(\lambda, \ell) \in \mathbb{V}^m \oplus \mathbb{V}^m$, is a **D-constraint torque-force** at Φ if $P(\lambda \oplus \ell, \psi)(0) = 0$ for every differentiable motion $\psi: [-r, r] \rightarrow \mathcal{Q}_{\text{free}}$ of the system satisfying $\psi(0) \in \mathcal{D}_\Phi$. We denote

$$\text{CT-F} = \{(\Phi, \tau \oplus f) \in \text{T-F} \mid (\tau, f) \text{ is a D-constraint torque-force at } \Phi\},$$

which is the **D-constraint force bundle** for the system. •

Let us explore the connection between our characterisation of constraint torque-forces and virtual displacements. There are some important differences between this and what we saw for interconnection torque-forces at the end of Section 6.6. First of all, a torque-force is a constraint torque-force exactly when the virtual power is zero for virtual displacements corresponding to infinitesimal D-variations. By parts (i) and (ii) of Proposition 6.20, such variations can be chosen to be Q-variations or not. However, by parts (iv) and (v) of Proposition 6.20, it is neither necessary nor sufficient to use virtual displacements corresponding to D-variations to characterise constraint torque-forces. Matters such as this are explored by Gràcia, Marin-Solano, and Muñoz-Lecanda [2003].

There is a large and sometimes vituperative literature around the subject of the preceding paragraph, and we shall overview a little of it, just for fun. First of all, we point out that, if one uses D-variations to characterise constraint torque-forces—and the variational principles associated to this—this gives the empirically incorrect governing equations. Nonetheless, this technique is often stated as being physically correct [e.g., Goldstein, Poole, Jr, and Safko 2001, §2.4], probably from some ill-placed belief in the primacy of variational principles. Indeed, this is the argument of Kozlov [1992] in support of this strategy of variations. This is argued against by Kharlomov [1992] in the same journal issue. Lewis and Murray [1995] show that the use of D-variations gives the physically correct equations of motion when the constraints are holonomic, i.e., when the constraints are actually on configurations and not on velocities (this is well-known and has been commented on by many authors). As well, Lewis and Murray give a specific example illustrating that there are physical motions that cannot arise from the use of D-variations. A somewhat recent overview of this subject is given by Flannery [2005]. There are many other “contributions” to this problem of variations, and it is probably one that will never really die. However, our view on this is that it is born partly from (1) the unnecessarily complicated and imprecise language in which the Principle of Virtual Work is couched and (2) an unjustified belief that Nature seeks to optimise in a certain restricted sense.⁵ The first of these problems is easily rectified by the following definition.⁶

6.23 Definition: (Principles of Virtual Power and Virtual Work) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Suppose that

⁵An appropriate minimisation, instantaneous in time, does give rise to the correct equations of motion. This is known as “Gauss’s Principle of Least Constraint,” and a modern presentation of this is given by Lewis [1996].

⁶The author is not able to rectify the second of these problems.

we have physical velocity constraints giving rise to the constraint distribution D . Let \mathbb{T} be a time axis and let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval.

- (i) The **Principle of Virtual Power** is: for every differentiable motion $\phi: \mathbb{T}' \rightarrow Q$ satisfying $\phi'(t) \in D_{\phi(t)}$,
 - (a) a continuous torque-force field $\tau \oplus \mathbf{f}$ along ϕ is an interconnection torque-force if and only if the virtual power of $\tau \oplus \mathbf{f}$ for every virtual displacement associated to an infinitesimal Q -variation of ϕ is zero;
 - (b) a continuous torque-force field $\tau \oplus \mathbf{f}$ along ϕ is a D -constraint torque-force if and only if the virtual power of $\tau \oplus \mathbf{f}$ for every virtual displacement associated to an infinitesimal D -variation of ϕ is zero.
- (ii) The **Principle of Virtual Work** is: for every differentiable motion $\phi: \mathbb{T}' \rightarrow Q$ satisfying $\phi'(t) \in D_{\phi(t)}$,
 - (a) a continuous torque-force field $\tau \oplus \mathbf{f}$ along ϕ for which $(\tau \oplus \mathbf{f}, \phi)$ is integrable is an interconnection torque-force if and only if the virtual work of $\tau \oplus \mathbf{f}$ for every virtual displacement associated to an infinitesimal Q -variation of ϕ is zero;
 - (b) a continuous torque-force field $\tau \oplus \mathbf{f}$ along ϕ for which $(\tau \oplus \mathbf{f}, \phi)$ is integrable is a D -constraint torque-force if and only if the virtual work of $\tau \oplus \mathbf{f}$ for every virtual displacement associated to an infinitesimal D -variation of ϕ is zero. •

Note that, in Definitions 6.21 and 6.22, we essentially use the Principle of Virtual Power along a trivial curve, i.e., a stationary curve defined on a time interval that is a point. This makes sense in this context since virtual power is an instantaneous notion.

We believe that our definitions give simple and physically meaningful statements that gives the correct equations of motion when they are translated into the problems of dynamics in Section 7. We know, however, that this will not be satisfying to some...

7. Interconnected rigid body systems: dynamics

In this section we present a clear account of the Newton–Euler equations for rigid body motion of single rigid bodies, interconnected systems of rigid bodies, and interconnected systems of rigid bodies subject to velocity constraints. We shall formulate the equations first for single rigid bodies, and then extend to multiple bodies with interconnections.

7.1. Momenta. The dynamical equations of mechanics are balance equations, and what is balanced is momentum in its appropriate forms. In this section we give the definitions of momentum, starting from the following basic physical constructions.

7.1 Lemma: (Translational and angular momentum) *Let $\mathcal{S} = (S, V, g, \theta)$ be a Newtonian space model, let $\mathcal{B} = (B, U, G, \Theta)$, and let (B, μ) be a rigid body in \mathcal{B} . Let $x_0 \in S$. Let \mathbb{T} be a time axis, let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval, and let $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ be a rigid*

motion. Then we have the following formulae:

$$\begin{aligned} \int_{\mathcal{B}} \frac{d}{dt} \phi(t)(X) d\mu(X) &= m \dot{r}_\phi(t); \\ \int_{\mathcal{B}} (\phi(t)(X) - x_0) \times \left(\frac{d}{dt} (\phi(t)(X) - x_0) \right) d\mu(X) &= m r_\phi(t) \times \dot{r}_\phi(t) + R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega(t)); \\ \int_{\mathcal{B}} (\phi(t)(X) - \phi(t)(X_c)) \times \left(\frac{d}{dt} (\phi(t)(X) - \phi(t)(X_c)) \right) d\mu(X) &= R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)). \end{aligned}$$

Proof: We have

$$\phi(t)(X) = x_0 + (r_\phi(t) + R_\phi(t)(X - X_c))$$

so

$$\frac{d}{dt} (\phi(t)(X)) = \dot{r}_\phi(t) + \dot{R}_\phi(t)(X - X_c).$$

Then the first formula follows immediately since

$$\int_{\mathcal{B}} (X - X_c) d\mu(X) = 0$$

by Lemma 3.4(ii). Now note that, by Lemma 2.17,

$$\hat{\omega}_\phi(t) = R_\phi(t) \circ \hat{\Omega}_\phi(t) \circ R_\phi^T(t) \implies \omega_\phi(t) = R_\phi(t)(\Omega_\phi(t)).$$

The second assertion of the lemma follows immediately from the preceding two facts and the computation

$$\begin{aligned} \int_{\mathcal{B}} (R_\phi(t)(X - X_c)) \times (R_\phi(t) \circ \hat{\Omega}_\phi(t)(X - X_c)) d\mu(X) \\ &= \int_{\mathcal{B}} R_\phi(t)((X - X_c) \times (\Omega_\phi(t) \times (X - X_c))) d\mu(X) \\ &= R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)). \end{aligned}$$

For the final assertion we have

$$\phi(t)(X - X_c) = R_\phi(t)(X - X_c)$$

and so compute

$$\begin{aligned} \int_{\mathcal{B}} (\phi(t)(X) - \phi(t)(X_c)) \times \left(\frac{d}{dt} (\phi(t)(X) - \phi(t)(X_c)) \right) d\mu(X) \\ &= \int_{\mathcal{B}} (R_\phi(t)(X - X_c)) \times (\dot{R}_\phi(t)(X - X_c)) d\mu(X) \\ &= \int_{\mathcal{B}} R_\phi(t)((X - X_c) \times (\hat{\Omega}_\phi(t)(X - X_c))) d\mu(X) \\ &= R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)), \end{aligned}$$

as desired. ■

Let us name the quantities in the lemma.

7.2 Definition: (Translational and angular momentum) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . Let $x_0 \in \mathbb{S}$. Let \mathbb{T} be a time axis, let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval, and let $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ be a rigid motion.

(i) The *spatial translational momentum* for the motion is

$$\begin{aligned} p_\phi: \mathbb{T}' &\rightarrow \mathbb{V} \\ t &\mapsto m\dot{r}_\phi(t). \end{aligned}$$

(ii) The *spatial angular momentum about x_0* for the motion is

$$\begin{aligned} \mu_{x_0, \phi}: \mathbb{T}' &\rightarrow \mathbb{V} \\ t &\mapsto m r_\phi(t) \times \dot{r}_\phi(t) + R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)). \end{aligned}$$

(iii) The *spatial angular momentum* for the motion is

$$\begin{aligned} \mu_\phi: \mathbb{T}' &\rightarrow \mathbb{V} \\ t &\mapsto R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)). \end{aligned} \quad \bullet$$

7.2. Newton–Euler equations for a single rigid body. Now we formulate the basic equations of rigid body dynamics, the “Newton–Euler equations.” In order to fully understand the origins of these equations, it is useful to start with a framework of distributed forces and torques.

7.3 Definition: (The integral form of the Newton–Euler equations) Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathbb{B}, μ) be a rigid body in \mathcal{B} . Let $x_0 \in \mathbb{S}$. Let \mathbb{T} be a time axis, let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval, and let $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ be a rigid motion. Let

$$\rho, \sigma: \mathbb{T}' \times \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \mathcal{M}(\mathbb{B}; \mathbb{V})$$

be weakly smooth torque distribution and force distribution fields. The motion ϕ satisfies the *integral form of Newton–Euler equations about x_0* for the torque distribution and force distribution fields ρ and σ if

$$\begin{aligned} \frac{d}{dt} \left(\int_{\mathbb{B}} \frac{d}{dt} \phi(t)(X) d\mu(X) \right) &= \int_{\mathbb{B}} d\sigma(t, \phi'(t))(X), \\ \frac{d}{dt} \left(\int_{\mathbb{B}} (\phi(t)(X) - x_0) \times \left(\frac{d}{dt} (\phi(t)(X) - x_0) \right) d\mu(X) \right) &= \\ &= \int_{\mathbb{B}} (\phi(t)(X) - x_0) \times d\sigma(t, \phi'(t))(X) + \int_{\mathbb{B}} d\rho(t, \phi'(t))(X) \end{aligned}$$

for every $t \in \mathbb{T}'$. •

By stating these equations in integral form, we see clearly how the mass distribution μ , and the torque distribution and force distribution fields ρ and σ allow one to state the equations of motion as infinitesimal (in space) momentum balance equations. However, these equations are clearly more conveniently expressed by doing the integrals and providing

“macroscopic” versions of the equations. In order to do this, associated with ρ and σ we define torque and force fields τ_ρ , τ_{σ, x_0} , and f_σ in the obvious way:

$$\begin{aligned}\tau_\rho(t, (\Phi, A, v)) &= \int_{\mathcal{B}} d\rho(t, (\Phi, (A, v)))(X); \\ \tau_{\sigma, x_0}(t, (\Phi, (A, v))) &= \int_{\mathcal{B}} (\Phi(X) - x_0) \times d\sigma(t, (\Phi, (A, v)))(X); \\ f_\sigma(t, (\Phi, (A, v))) &= \int_{\mathcal{B}} d\sigma(t, (\Phi, (A, v)))(X).\end{aligned}$$

With these constructions we can state the following result which lists various equivalent forms of the Newton–Euler equations.

7.4 Theorem: (Macroscopic forms of the Newton–Euler equations) *Let $\mathcal{S} = (\mathcal{S}, \mathcal{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathcal{B}, \mathcal{U}, \mathcal{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . Let $x_0 \in \mathcal{S}$. Let \mathbb{T} be a time axis and let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval. Let*

$$\rho, \sigma: \mathbb{T}' \times \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) \rightarrow \mathcal{M}(\mathcal{B}; \mathcal{V})$$

be weakly smooth torque distribution and force distribution fields. For a rigid motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$, the following statements are equivalent:

- (i) ϕ satisfies the integral form of the Newton–Euler equations about x_0 for the torque distribution and force distribution fields ρ and σ ;
- (ii) ϕ satisfies the equations

$$\begin{aligned}\dot{p}_\phi(t) &= f_\sigma(t, \phi'(t)), \\ \dot{\mu}_{x_0, \phi}(t) &= \tau_{\sigma, x_0}(t, \phi'(t)) + \tau_\rho(t, \phi'(t))\end{aligned}$$

for every $t \in \mathbb{T}'$;

- (iii) ϕ satisfies the equations

$$\begin{aligned}\dot{p}_\phi(t) &= f_\sigma(t, \phi'(t)), \\ \dot{\mu}_\phi(t) &= \tau_{\sigma, \phi(t)(X_c)}(t, \phi'(t)) + \tau_\rho(t, \phi'(t))\end{aligned}$$

for every $t \in \mathbb{T}'$;

- (iv) ϕ satisfies the equations

$$\begin{aligned}m\ddot{r}_\phi(t) &= f_\sigma(t, \phi'(t)), \\ R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\dot{\omega}_\phi(t)) + \omega_\phi(t) \times (R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t))) &= \\ &= \tau_{\sigma, \phi(t)(X_c)}(t, \phi'(t)) + \tau_\rho(t, \phi'(t))\end{aligned}$$

for every $t \in \mathbb{T}'$;

- (v) ϕ satisfies the equations

$$\begin{aligned}m\dot{V}_\phi(t) + \Omega_\phi(t) \times (mV_\phi(t)) &= R_\phi^T(t)(f_\sigma(t, \phi'(t))) \\ \mathbb{I}_c(\dot{\Omega}_\phi(t)) + \Omega_\phi(t) \times (\mathbb{I}_c(\Omega_\phi(t))) &= R_\phi^T(t)(\tau_{\sigma, \phi(t)(X_c)}(t, \phi'(t)) + \tau_\rho(t, \phi'(t)))\end{aligned}$$

for every $t \in \mathbb{T}'$.

Proof: The equivalence of parts (i) and (ii) follows immediately just by the definition of the symbols and Definitions 7.2 and 7.3.

(ii) \iff (iii) If either of the equations of parts (ii) or (iii) hold, then $m\ddot{r}_\phi(t) = f_\sigma(t, \phi'(t))$. Given this, we then calculate

$$\dot{\mu}_{x_0, \phi}(t) - \dot{\mu}_\phi(t) = r_\phi(t) \times (m\ddot{r}_\phi(t)) = r_\phi(t) \times f_\sigma(t, \phi'(t)).$$

We also calculate

$$\tau_{\sigma, x_0}(t, (\Phi, (A, v))) - \tau_{\sigma, \Phi(X_c)}(t, (\Phi, (A, v))) = \int_{\mathcal{B}} (\Phi(X_c) - x_0) \times d\sigma(t, (\Phi, (A, v)))(X)$$

which gives

$$\tau_{\sigma, x_0}(t, \phi'(t)) - \tau_{\sigma, \phi(t)(X_c)}(t, \phi'(t)) = r_\phi(t) \times f_\sigma(t, \phi'(t)).$$

Combining the preceding two computations gives the desired equivalence.

Let us now prove the equivalence of the final three statements of the theorem. For brevity, let us denote

$$\bar{\mathbb{I}}_c(t) = R_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t).$$

We first compute

$$\begin{aligned} \dot{\mu}_\phi(t) &= \dot{R}_\phi(t) \circ \mathbb{I}_c \circ R_\phi^T(t)(\omega_\phi(t)) + R_\phi(t) \circ \mathbb{I}_c \circ \dot{R}_\phi^T(t)(\omega_\phi(t)) + \bar{\mathbb{I}}_c(t)(\dot{\omega}_\phi(t)) \\ &= \dot{R}_\phi(t) \circ R_\phi^T(t) \circ \bar{\mathbb{I}}_c(t)(\omega_\phi(t)) + \bar{\mathbb{I}}_c(t) \circ R_\phi(t) \circ \dot{R}_\phi^T(t)(\omega_\phi(t)) + \bar{\mathbb{I}}_c(t)(\dot{\omega}_\phi(t)) \\ &= \hat{\omega}_\phi(t) \circ \bar{\mathbb{I}}_c(t)(\omega_\phi(t)) + \bar{\mathbb{I}}_c(t) \circ \hat{\omega}_\phi(t)(\omega_\phi(t)) + \bar{\mathbb{I}}_c(t)(\dot{\omega}_\phi(t)) \\ &= \bar{\mathbb{I}}_c(t)(\dot{\omega}_\phi(t)) + \omega_\phi(t) \times (\bar{\mathbb{I}}_c(t)(\omega_\phi(t))). \end{aligned}$$

We will use this computation a few times below.

(iii) \implies (iv) This follows immediately from the preceding calculation.

(iv) \implies (v) First we compute

$$\begin{aligned} m\ddot{r}_\phi(t) &= \dot{p}_\phi(t) = m \frac{d}{dt}(R_\phi(t)V_\phi(t)) \\ &= m\dot{R}_\phi(t)(V_\phi(t)) + mR_\phi(t)(\dot{V}_\phi(t)) \\ &= R_\phi(t)(m(\hat{\Omega}_\phi(t)(V_\phi(t)) + \dot{V}_\phi(t))), \end{aligned}$$

which gives

$$m\ddot{r}_\phi(t) = f(t, \phi'(t)) \iff m\dot{V}_\phi(t) + \Omega_\phi(t) \times (mV_\phi(t)) = R_\phi^T(t)(f(t, \phi'(t))). \quad (7.1)$$

We also have, using Lemma 2.17,

$$\begin{aligned} \dot{\mu}_\phi(t) &= \frac{d}{dt}(R_\phi(t) \circ \mathbb{I}_c(\Omega_\phi(t))) \\ &= \dot{R}_\phi(t) \circ \mathbb{I}_c(\Omega_\phi(t)) + R_\phi(t) \circ \mathbb{I}_c(\dot{\Omega}_\phi(t)) \\ &= R_\phi(t)(\hat{\Omega}_\phi(t)(\mathbb{I}_c(\Omega_\phi(t))) + \mathbb{I}_c(\dot{\Omega}_\phi(t))), \end{aligned}$$

which gives

$$\dot{\mu}_\phi(t) = \tau(t, \phi'(t)) \iff \mathbb{I}_c(\dot{\Omega}_\phi(t)) + \Omega_\phi(t) \times (\mathbb{I}_c(\Omega_\phi(t))) = R_\phi^T(t)(\tau(t, \phi'(t))). \quad (7.2)$$

Taking

$$\tau(t, \phi'(t)) = \tau_{\sigma, \phi(t)(X_c)}(t, \phi'(t)) + \tau_\rho(t, \phi'(t)),$$

this part of the theorem then follows from the computation preceding the proof of the implication (iii) \implies (iv).

(v) \implies (iii) This follows since the implications of equations (7.1) and (7.2) go both ways. \blacksquare

The equations of part (ii) are the *spatial Newton–Euler equations about x_0* and the equations of part (iii) are the *spatial Newton–Euler equations*. The equivalence of these two forms of the Newton–Euler equations is a reflection of the well-known principle that one can obtain physically correct equations of motion by either (1) balancing moments about a spatially fixed point or (2) balancing moments about the centre of mass of a body. The equations of part (iv) are simply a re-expression of the spatial Newton–Euler equations with velocity derivatives instead of momentum derivatives. Thus these equations are to be regarded as spatial, as with the first three parts of the theorem. The only “body” equations are those of part (v), which we call the *body Newton–Euler equations*.

7.3. Newton–Euler equations for interconnected rigid body systems with velocity constraints. Now we turn to the Newton–Euler equations for multiple rigid bodies, allowing for interconnections and velocity constraints. By virtue of Theorem 7.4, we can consider central forces and torques by taking suitable resultant torques for non-central force distributions. Therefore, without loss of generality for the dynamics, we can work only with force and torque fields in the usual sense.

The following definition gives the required construction.

7.5 Definition: (Newton–Euler equations for interconnected rigid body systems with velocity constraints) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathbb{Q} . Let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. Let $\tau \oplus \mathbf{f}$ be a torque-force field and suppose that we have velocity constraints giving rise to a smooth locally constant rank constraint distribution \mathbb{D} . A C^2 -motion $\phi: \mathbb{T}' \rightarrow \mathbb{Q}$ for which $\phi'(t) \in \mathbb{D}_{\phi(t)}$ satisfies the *Newton–Euler equations* if there exist

- (i) an IT-F-valued torque-force field $\iota \oplus \mathbf{i}: \mathbb{T}' \rightarrow \mathbb{V}^m \oplus \mathbb{V}^m$ along ϕ and
- (ii) a CT-F-valued torque-force field $\lambda \oplus \mathbf{\ell}: \mathbb{T}' \rightarrow \mathbb{V}^m \oplus \mathbb{V}^m$ along ϕ

such that

$$\begin{aligned} \dot{p}_{\phi_a}(t) &= f_a(t, \phi'(t)) + i_a(t) + \ell_a(t), \\ \dot{m}_{\phi_a}(t) &= \tau_a(t, \phi'(t)) + \iota_a(t) + \lambda_a(t) \end{aligned}$$

for each $a \in \{1, \dots, m\}$ and $t \in \mathbb{T}'$. \bullet

The Newton–Euler equations are, in essence, differential-algebraic equations. As such, the matter of existence and uniqueness of solutions is not just immediately clear. However, we shall see in Theorem 8.18, when the data is sufficiently smooth, these equations are actually equivalent to well-posed second-order differential equations (the constrained Euler–Lagrange equations, of course), and so the matter of existence and uniqueness of solutions becomes that for standard ordinary differential equations.

7.4. Lagrange–d’Alembert Principle. The Newton–Euler equations can alternatively be framed using the Lagrange–d’Alembert Principle. For our purposes, this will provide a convenient device for the transition from the balance framework of the Newton–Euler equations to the variational framework of the Euler–Lagrange equations.

In our setting, the following result gives what we want.

7.6 Proposition: (Lagrange–d’Alembert Principle (instantaneous Newton–Euler version)) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Let $\boldsymbol{\tau} \oplus \mathbf{f}$ be a torque-force field and suppose that we have velocity constraints giving rise to a smooth locally constant rank constraint distribution \mathcal{D} . Let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. For a C^2 -motion $\phi: \mathbb{T}' \rightarrow \mathcal{Q}$ satisfying $\phi'(t) \in \mathcal{D}_{\phi(t)}$, the following statements are equivalent:

- (i) ϕ satisfies the Newton–Euler equations;
- (ii) ϕ satisfies the **instantaneous Lagrange–d’Alembert Principle**: for every $t \in \mathbb{T}'$,

$$\sum_{a=1}^m (\mathfrak{g}(\dot{\mu}_{\phi_a}(t) - \tau_a(t, \phi'(t)), \omega_{\psi_a}(0)) + \mathfrak{g}(\dot{p}_{\phi_a}(t) - f_a(t, \phi'(t)), v_{\psi_a}(0))) = 0$$

for every differentiable motion $\psi: [-r, r] \rightarrow \mathcal{Q}_{\text{free}}$ for which $\psi'(0) \in \mathcal{D}_{\phi(t)}$.

Proof: (i) \implies (ii) If ϕ is a solution to the Newton–Euler equations, then there exist torque-force fields $\boldsymbol{\iota} \oplus \mathbf{i}$ and $\boldsymbol{\lambda} \oplus \boldsymbol{\ell}$ along ϕ such that:

1. for every $t \in \mathbb{T}'$,

$$\sum_{a=1}^m (\mathfrak{g}(\iota_a(t), \omega_{\psi_a}(0)) + \mathfrak{g}(i_a(t), v_{\psi_a}(0))) = 0$$

for every differentiable motion $\psi: [-r, r] \rightarrow \mathcal{Q}$ for which $\psi(0) = \phi(t)$;

2. for every $t \in \mathbb{T}'$,

$$\sum_{a=1}^m (\mathfrak{g}(\lambda_a(t), \omega_{\psi_a}(0)) + \mathfrak{g}(\ell_a(t), v_{\psi_a}(0))) = 0$$

for every differentiable motion $\psi: [-r, r] \rightarrow \mathcal{Q}$ for which $\psi'(0) \in \mathcal{D}_{\phi(t)}$;

3. we have

$$\begin{aligned} \dot{p}_{\phi_a}(t) - f_a(t, \phi'(t)) &= i_a(t) + \ell_a(t), \\ \dot{\mu}_{\phi_a}(t) - \tau_a(t, \phi'(t)) &= \iota_a(t) + \lambda_a(t) \end{aligned}$$

for each $a \in \{1, \dots, m\}$ and $t \in \mathbb{T}'$.

This immediately implies that ϕ satisfies the instantaneous Lagrange–d’Alembert Principle.

(ii) \implies (i) Suppose that ϕ satisfies the instantaneous Lagrange–d’Alembert Principle and define a torque-force field $\mathbf{T} \oplus \mathbf{F}$ along ϕ by

$$T_a(t) = \dot{\mu}_{\phi_a}(t) - \tau_a(t, \phi'(t)), \quad F_a(t) = \dot{p}_{\phi_a}(t) - f_a(t, \phi'(t)),$$

$t \in \mathbb{T}'$, $a \in \{1, \dots, m\}$. Note that $\mathbf{T} \oplus \mathbf{F}$ is already CT-F-valued, and so we can conclude the proof by taking $\boldsymbol{\iota} \oplus \mathbf{i} = 0$ and $\boldsymbol{\lambda} \oplus \boldsymbol{\ell} = \mathbf{T} \oplus \mathbf{F}$. However, many other choices are possible. For example, we can require that the constraint force satisfy

$$\boldsymbol{\lambda} \oplus \boldsymbol{\ell}(t) \in \boldsymbol{\tau}_{\text{spatial}}(\mathbf{TQ}),$$

and then take

$$\boldsymbol{\iota} \oplus \mathbf{i}(t) = \mathbf{T} \oplus \mathbf{F}(t) - \boldsymbol{\lambda} \oplus \boldsymbol{\ell}(t),$$

which conditions uniquely prescribe the interconnection and constraint force-torques, and as well has the property that

$$\sum_{a=1}^m (\mathbf{g}(\boldsymbol{\iota}_a(t), \boldsymbol{\lambda}_a(t)) + \mathbf{g}(\mathbf{i}_a(t), \boldsymbol{\ell}_a(t))) = 0,$$

i.e., the interconnection and constraint torque-forces are orthogonal in an appropriate sense. ■

7.7 Remarks: (On the Lagrange–d’Alembert Principle)

1. The proof of the preceding proposition bears out the fact that, while there is no ambiguity in the assignment of $(\boldsymbol{\iota} + \boldsymbol{\lambda}) \oplus (\mathbf{i} + \boldsymbol{\ell})$, each of $\boldsymbol{\iota} \oplus \mathbf{i}$ and $\boldsymbol{\lambda} \oplus \boldsymbol{\ell}$ are not uniquely defined by the the requirement that they determine a solution to the Newton–Euler equations.
2. Note that, for a C^2 -motion $\phi: \mathbb{T}' \rightarrow \mathbf{Q}$, we have a torque-force field $\boldsymbol{\tau}_\phi \oplus \mathbf{f}_\phi$ along ϕ defined by

$$\boldsymbol{\tau}_{\phi,a}(t) = \dot{\boldsymbol{\mu}}_{\phi_a}(t), \quad \mathbf{f}_{\phi,a}(t) = \dot{\mathbf{p}}_{\phi_a}(t),$$

$t \in \mathbb{T}'$, $a \in \{1, \dots, m\}$. This we can call the *inertial torque-force field* along the motion. Note that, for the a th body and for $t \in \mathbb{T}'$, $\boldsymbol{\tau}_{\phi,a}(t) \oplus \mathbf{f}_{\phi,a}(t)$ is the central torque-force associated with the inertial force and inertial torque distribution of Example 6.2–4 and Example 6.8–3.

The instantaneous Lagrange–d’Alembert Principle is then exactly the Principle of Virtual Power applied to the difference between the inertial and applied torque-force fields along the motion. When \mathbb{T}' is compact (to ensure integrals exist), this is equivalent to

$$\sum_{a=1}^m \int_{\mathbb{T}'} \mathbf{g}(\dot{\boldsymbol{\mu}}_{\phi_a}(t) - \boldsymbol{\tau}_a(t, \phi'(t)), \boldsymbol{\omega}_{\sigma_a}(t)) dt + \sum_{a=1}^m \int_{\mathbb{T}'} \mathbf{g}(\dot{\mathbf{p}}_{\phi_a}(t) - \mathbf{f}_a(t, \phi'(t)), \mathbf{v}_{\sigma_a}(t)) dt = 0$$

for every infinitesimal D-variation $\boldsymbol{\sigma}$ of ϕ (we refer to Section 6.5 for notation). This is the typical form in which one sees the Lagrange–d’Alembert Principle presented. However, the advantage of using the integral formulation is not presently clear, and will only be seen in Section 8 when we connect dynamics to variational principles. •

7.5. Kinetic energy. A crucial ingredient in connecting the Newton–Euler equations in this section to the Euler–Lagrange equations of Section 8 is the kinetic energy Lagrangian. In this section we show how kinetic energy is connected to the Riemannian (semi)metrics of Section 5.4.

We start by considering a single rigid body.

7.8 Lemma: (Kinetic energy of a rigid body) *Let $\mathcal{S} = (\mathbf{S}, \mathbf{V}, \mathbf{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbf{B}, \mathbf{U}, \mathbf{G}, \Theta)$ be a body reference space, let (\mathcal{B}, μ) be a rigid body in \mathcal{B} , let \mathbb{T} be a time axis, let $\mathbb{T}' \subseteq \mathbb{T}$ be a time interval, and let $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ be a rigid motion. Then*

$$\frac{1}{2} \int_{\mathcal{B}} \mathbf{g}\left(\frac{\text{d}}{\text{d}t}(\phi(t)(X)), \frac{\text{d}}{\text{d}t}(\phi(t)(X))\right) \text{d}\mu(X) = \frac{1}{2} m \mathbf{g}(\dot{r}_\phi(t), \dot{r}_\phi(t)) + \frac{1}{2} \mathbf{G}(\mathbf{I}_c(\Omega_\phi(t)), \Omega_\phi(t)).$$

Proof: Let $x_0 \in \mathbf{S}$; as we shall see, the argument does not depend on the selection of such a spatial origin. We have

$$\phi(t)(X) = x_0 + (r_\phi(t) + R_\phi(t)(X - X_c)) \implies \frac{\text{d}}{\text{d}t}\phi(t)(X) = \dot{r}_\phi(t) + \dot{R}_\phi(t)(X - X_c).$$

Thus

$$\begin{aligned} & \mathbf{g}\left(\frac{\text{d}}{\text{d}t}(\phi(t)(X)), \frac{\text{d}}{\text{d}t}(\phi(t)(X))\right) \\ &= \mathbf{g}(\dot{r}_\phi(t), \dot{r}_\phi(t)) + 2\mathbf{g}(\dot{r}_\phi(t), \dot{R}_\phi(t)(X - X_c)) + \mathbf{G}(\hat{\Omega}_\phi(t)(X - X_c), \hat{\Omega}_\phi(t)(X - X_c)). \end{aligned}$$

Note that

$$\int_{\mathcal{B}} \mathbf{g}(\dot{r}_\phi(t), \dot{R}_\phi(t)(X - X_c)) \text{d}\mu(X) = \mathbf{g}\left(\dot{r}_\phi(t), \dot{R}_\phi(t) \int_{\mathcal{B}} (X - X_c) \text{d}\mu(X)\right) = 0$$

by Lemma 3.4(ii). Note that

$$\mathbf{G}(\hat{\Omega}_\phi(t)(X - X_c), \hat{\Omega}_\phi(t)(X - X_c)) = \mathbf{G}(\Omega_\phi(t), (X - X_c) \times (\Omega_\phi(t) \times (X - X_c))),$$

using (1.1). By definition of the inertia tensor, the result follows. \blacksquare

Correspondingly with our integral characterisations of momentum and the Newton–Euler equations, the preceding lemma gives the characterisation of kinetic energy as being an infinitesimal version of “ $\frac{1}{2}mv^2$,” integrated over the body.

Next we provide the multi-body version of the preceding lemma, incorporating interconnections and degenerate rigid bodies. In (5.2) we defined a Riemannian semimetric $\mathbf{G}_{\mathcal{B}}$ on \mathbf{Q}_{free} . This drops to the Riemannian metric $\mathbf{G}_{0, \mathcal{B}}$ on the reduced free configuration manifold $\mathbf{Q}_{0, \text{free}}$ defined by (5.3). The following result connects these constructions with the kinetic energy of the interconnected system.

7.9 Lemma: (Kinetic energy and Riemannian metrics) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and with configuration manifold \mathbf{Q} . Let \mathbb{T} be a time axis and $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. Then, for any motion $\phi: \mathbb{T}' \rightarrow \mathbf{Q}$ of the system, we have*

$$\begin{aligned} \sum_{a=1}^m \frac{1}{2} \int_{\mathcal{B}_a} \mathbf{g}\left(\frac{\text{d}}{\text{d}t}(\phi_a(t)(X_a)), \frac{\text{d}}{\text{d}t}(\phi_a(t)(X_a))\right) \text{d}\mu_a(X_a) &= \frac{1}{2} \mathbf{G}(\phi'(t), \phi'(t)) \\ &= \frac{1}{2} \mathbf{G}_0(\phi'_0(t), \phi'_0(t)), \quad t \in \mathbb{T}'. \end{aligned}$$

Proof: The first equality is an immediate consequence of the definition of $\mathbf{G}_{\mathcal{B}}$ in equation (5.2) and from Definition 3.7, and noting that \mathbf{G} is just the restriction of $\mathbf{G}_{\mathcal{B}}$ to \mathbf{Q} . The second equality is an immediate consequence of Lemma 4.17, again along with the fact that \mathbf{G}_0 is just the restriction of $\mathbf{G}_{0, \mathcal{B}}$ to \mathbf{Q}_0 . \blacksquare

8. Interconnected rigid body systems: Euler–Lagrange equations

In this section we develop the Euler–Lagrange equations from the Newton–Euler equations of Section 7. Let us give the “quick and dirty” version of our transition from Newton–Euler to Euler–Lagrange in the case when all rigid bodies of the system have nondegenerate inertia tensors. First we establish, using variational arguments following Marsden and Scheurle [1993], the equivalence of the unforced Newton–Euler equations with geodesic equations on $\text{Rgd}(\mathcal{B}; \mathcal{S})$ (we refer to the introduction of [Cendra, Holm, Marsden, and Ratiu 1998] for an historical account of the development of this approach). If one has multiple bodies, then their free, i.e., non-interconnected, motion is simply a product of these geodesic equations. In order to consider external forces, interconnections, and velocity constraints, we develop how a torque-force is translated into a cotangent vector on the configuration manifold of the system. We then use the Lagrange–d’Alembert Principle in various guises to prove the equivalence of the general Newton–Euler equations to the forced Euler–Lagrange equations. A lot of the bulk in our presentation is a result of the fact that we *do* consider degenerate rigid bodies. It is worth pointing out, however, that all of our work with degenerate rigid bodies comes together in this section, as *all* of the work we have done for such bodies over Sections 4–7 gets used here.

The tools for the approach are developed in Sections 8.1, 8.2, and 8.3.

8.1. Some constructions in variational calculus. In this section we review the elements of variational calculus needed for our arc from Newton–Euler to Euler–Lagrange. Variational calculus is set up as a means of tackling certain sorts of optimisation problems, typically optimisation problems over spaces of curves in its simplest guise. We, however, are not concerned with the details of this connection to optimisation problems, but are only concerned with the first-order conditions for optimisation, i.e., first derivative conditions. These first-order conditions, as is well-known, give rise to the Euler–Lagrange equations for an optimisation problem, and solutions to these equations are called “extremals.” Thus our discussion of variational calculus is limited to the definition of extremals. Moreover, since we are only interested in differential equations, which by nature are local, we can make simplifying assumptions to the setup that will keep us from having to fiddle with unnecessary technicalities, while maintaining mathematical rigour. Our presentation in this section is general, so we work with an arbitrary smooth paracompact Hausdorff manifold M .

First let us consider the sort of curves we will work with.

8.1 Definition: (Arc) A C^2 -arc on a manifold M is a C^2 -mapping $\gamma: [t_0, t_1] \rightarrow M$. •

Now we can vary arcs.

8.2 Definition: (Fixed endpoint variation of arc, infinitesimal variation) Let $\gamma: [t_0, t_1] \rightarrow M$ be a C^2 -arc. A *fixed endpoint variation* of γ is a C^2 -mapping $\sigma: [-r, r] \times [t_0, t_1] \rightarrow M$, $r \in \mathbb{R}_{>0}$, such that

- (i) $\sigma(0, t) = \gamma(t)$, $t \in [t_0, t_1]$,
- (ii) $\sigma(s, t_0) = \gamma(t_0)$ and $\sigma(s, t_1) = \gamma(t_1)$, $s \in [-r, r]$.

The *infinitesimal variation* of a fixed endpoint variation σ is the mapping

$$\begin{aligned} \delta\sigma: [t_0, t_1] &\rightarrow \text{TM} \\ t &\mapsto \left. \frac{d}{ds} \right|_{s=0} \sigma(s, t). \end{aligned} \quad \bullet$$

Of course we have seen these ideas before when discussing virtual displacements in Section 6.5. Here we are working in a non-physical setting, and so prefer to use non-physical language.

For $s \in [-r, r]$, let us denote $\sigma_s: [t_0, t_1] \rightarrow \text{M}$ by $\sigma_s(t) = \sigma(s, t)$. We then denote $\sigma'(s, t) \triangleq \sigma'_s(t)$.

Let us show that sufficiently rich classes of variations exist.

8.3 Lemma: (Existence of variations with prescribed infinitesimal variations) *Let $\gamma: [t_0, t_1] \rightarrow \text{M}$ be a C^2 -arc and let $V: [t_0, t_1] \rightarrow \text{TM}$ be a C^1 -map with the following properties:*

- (i) $V(t) \in \mathbb{T}_{\gamma(t)}\text{M}$;
- (ii) $V(t_0) = 0_{\gamma(t_0)}$ and $V(t_1) = 0_{\gamma(t_1)}$.

Then there exists $r \in \mathbb{R}_{>0}$ and a fixed endpoint variation $\sigma: [-r, r] \times [t_0, t_1] \rightarrow \text{M}$ such that $V(t) = \delta\sigma(t)$ for all $t \in [t_0, t_1]$.

Proof: Let \mathbf{G} be a Riemannian metric on M and, for $x \in \text{M}$, let \exp_x be the exponential map which is defined on a neighbourhood of $0_x \in \mathbb{T}_x\text{M}$ taking values in M . By a standard compactness argument, there exists $r \in \mathbb{R}_{>0}$ such that $\exp_x(sV(t))$ is defined for each $(s, t) \in [-r, r] \times [t_0, t_1]$. We claim that the map

$$\begin{aligned} \sigma: [-r, r] \times [t_0, t_1] &\rightarrow \text{M} \\ (s, t) &\mapsto \exp_{\gamma(t)}(sV(t)) \end{aligned}$$

is of class C^2 . Indeed, we note that σ is the composition of the mappings

$$\begin{aligned} \bar{V}: [-r, r] \times [t_0, t_1] &\rightarrow [-r, r] \times \text{TM} \\ (s, t) &\mapsto (s, V(t)) \end{aligned}$$

and

$$\begin{aligned} F: [-r, r] \times \text{TM} &\rightarrow \text{M} \\ (s, v_x) &\mapsto \exp_x(sv_x). \end{aligned}$$

The first mapping is of class C^2 . The second is smooth since it is the time 1 flow of the vector field $\frac{\partial}{\partial s} \oplus Z$ on $[-r, r] \times \text{TM}$, Z being the geodesic spray for the Levi-Civita affine connection associated with \mathbf{G} . Moreover, since $s \mapsto \exp_x(sv_x)$ is the geodesic with initial velocity v_x , we have $\delta\sigma(t) = V(t)$ for every $t \in [t_0, t_1]$. \blacksquare

Now we consider extremal arcs associated to a Lagrangian.

8.4 Definition: (Lagrangian, extremal) For a manifold M :

- (i) a *Lagrangian* is a smooth function L on TM ;
- (ii) an *L-extremal* is a C^2 -arc $\gamma: [t_0, t_1] \rightarrow M$ such that

$$\delta_\sigma L \triangleq \left. \frac{d}{ds} \right|_{s=0} \int_{t_0}^{t_1} L(\sigma'(s, t)) dt = 0$$

for every fixed endpoint variation σ of γ . •

As mentioned in the introductory remarks to this section, while L extremals are connected to the minimisation of the function

$$\gamma \mapsto \int_{t_0}^{t_1} L(\gamma'(t)) dt,$$

we shall not make use of this fact. We shall also not write the familiar coordinate expression for the Euler–Lagrange equations [Goldstein, Poole, Jr, and Safko 2001, Equation (2.18)] for extremals as this would violate our “no coordinates” rule for the paper. We shall be able to write intrinsic forms of these equations in the cases of interest. Nonetheless, in order to connect our language with standard terminology, we can say that a curve γ is a *solution to the Euler–Lagrange equations* for the Lagrangian L if it is an L -extremal.

8.2. Affine connections, distributions, and submanifolds. In the next section we will consider an abstract framework for mechanics, and in doing so we will get some benefit from seeing how affine connections, particularly the Levi-Civita connection associated with a Riemannian metric, interact with distributions and submanifolds. We give here a self-contained treatment of the material we require from the paper of Lewis [1998].

We first work with a Riemannian manifold (M, \mathbb{G}) with a smooth, locally constant rank distribution $D \subseteq TM$. We have well-defined orthogonal projections

$$P_D: TM \rightarrow TM, \quad P_D^\perp: TM \rightarrow TM$$

onto D and $D^{\perp\mathbb{G}}$, respectively. We then have the affine connection $\overset{D}{\nabla}$ on M given by

$$\overset{D}{\nabla}_X Y = \overset{\mathbb{G}}{\nabla}_X Y + (\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y),$$

thinking of $\overset{\mathbb{G}}{\nabla}_X P_D^\perp$ as a $(1, 1)$ -tensor field on M , i.e., a section of the endomorphism bundle of TM . The following lemma provides some attributes of these constructions.

8.5 Lemma: (Affine connections and distributions) *The affine connection $\overset{D}{\nabla}$ has the following properties:*

- (i) $\overset{D}{\nabla}_X Y$ is D -valued for every D -valued vector field Y ;
- (ii) $\overset{D}{\nabla}_X Y - \overset{\mathbb{G}}{\nabla}_X Y$ is $D^{\perp\mathbb{G}}$ -valued for every D -valued vector field Y .

Proof: (i) For vector fields X and Y we have

$$P_D^\perp(\overset{D}{\nabla}_X Y) = P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y) + P_D^\perp(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y).$$

For a D -valued vector field Y and any vector field X we have

$$\begin{aligned} P_D^\perp(Y) &= 0 \\ \implies (\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) + P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y) &= 0 \\ \implies P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y) + P_D^\perp(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) &= 0 \end{aligned}$$

since $P_D^\perp \circ P_D^\perp = P_D^\perp$. Combing the preceding two calculations gives $P_D^\perp(\overset{D}{\nabla}_X Y) = 0$ if Y is D -valued, as desired.

(ii) If Y is D -valued, as above we have

$$(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) + P_D^\perp(\overset{\mathbb{G}}{\nabla}_X Y) = 0,$$

which gives

$$P_D((\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y)) = 0$$

since $P_D \circ P_D^\perp = 0$. This gives the result. \blacksquare

We next apply this construction in a special case.

8.6 Lemma: (Affine connections and submanifolds) *Let (M, \mathbb{G}) be a Riemannian manifold. If $D \subseteq TM$ is a smooth integrable distribution and if $M_0 \subseteq M$ is an embedded submanifold that is an integral manifold for D , then $\overset{D}{\nabla}$ restricts to an affine connection on M_0 that is the Levi-Civita connection associated to the Riemannian metric \mathbb{G}_0 obtained by restricting \mathbb{G} .*

Proof: First of all, by Lemma 8.5(i) follows the assertion that $\overset{D}{\nabla}$ restricts to an affine connection on M_0 . We shall denote this affine connection just by $\overset{D}{\nabla}$. Smooth vector fields on M_0 can be locally extended to smooth vector fields on M since M_0 is an embedded submanifold. We shall make such extensions without mention.

We must show that $\overset{D}{\nabla}$ is torsion-free and preserves the metric \mathbb{G}_0 . We first show that $\overset{D}{\nabla}$ is torsion-free. Let X and Y be vector fields on M_0 . We have

$$\overset{D}{\nabla}_X Y - \overset{D}{\nabla}_Y X = \overset{\mathbb{G}}{\nabla}_X Y - \overset{\mathbb{G}}{\nabla}_Y X + (\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) - (\overset{\mathbb{G}}{\nabla}_Y P_D^\perp)(X). \quad (8.1)$$

By Lemma 8.5 we have that $(\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y)$ and $(\overset{\mathbb{G}}{\nabla}_Y P_D^\perp)(X)$ are \mathbb{G} -orthogonal to TM_0 and that $\overset{D}{\nabla}_X Y$ and $\overset{D}{\nabla}_Y X$ are TM_0 -valued. Since $\overset{D}{\nabla}$ is torsion-free and since D is integrable,

$$\overset{\mathbb{G}}{\nabla}_X Y - \overset{\mathbb{G}}{\nabla}_Y X = [X, Y]$$

is D -valued if X and Y are D -valued. Matching the D - and D^\perp -components of (8.1), we obtain

$$\overset{D}{\nabla}_X Y - \overset{D}{\nabla}_Y X = [X, Y], \quad (\overset{\mathbb{G}}{\nabla}_X P_D^\perp)(Y) = (\overset{\mathbb{G}}{\nabla}_Y P_D^\perp)(X),$$

which shows, in particular, that $\overset{D}{\nabla}$ is torsion-free.

Now we show that $\overset{\text{D}}{\nabla}$, restricted to M_0 , preserves \mathbf{G}_0 . Let X, Y , and Z be vector fields on M_0 . Since $\overset{\text{G}}{\nabla}$ is the Levi-Civita connection,

$$\mathcal{L}_X(\mathbf{G}(Y, Z)) = \mathbf{G}(\overset{\text{G}}{\nabla}_X Y, Z) + \mathbf{G}(Y, \overset{\text{G}}{\nabla}_X Z).$$

Now we have

$$\mathbf{G}(\overset{\text{G}}{\nabla}_X Y, Z) = \mathbf{G}(\overset{\text{D}}{\nabla}_X Y, Z) - \mathbf{G}((\overset{\text{G}}{\nabla}_X P_{\text{D}}^\perp)(Y), Z) = \mathbf{G}(\overset{\text{D}}{\nabla}_X Y, Z)$$

since $(\overset{\text{G}}{\nabla}_X P_{\text{D}}^\perp)(Y)$ is \mathbf{G} -orthogonal to D by Lemma 8.5. Similarly, we may show that

$$\mathbf{G}(Y, \overset{\text{G}}{\nabla}_X Z) = \mathbf{G}(Y, \overset{\text{D}}{\nabla}_X Z).$$

Therefore,

$$\mathcal{L}_X(\mathbf{G}(Y, Z)) = \mathbf{G}(\overset{\text{D}}{\nabla}_X Y, Z) + \mathbf{G}(Y, \overset{\text{D}}{\nabla}_X Z)$$

which, when restricted to M_0 , shows that $\overset{\text{D}}{\nabla}$ preserves \mathbf{G}_0 . ■

8.3. Mechanics on abstract Riemannian manifolds. In this section we consider a Riemannian manifold (M, \mathbf{G}) and make mechanics-like constructions on it that we will subsequently use to connect the Newton–Euler equations to the Euler–Lagrange equations.

The first construction we make provides a characterisation of geodesics as extremals of a certain Lagrangian, namely the “kinetic energy Lagrangian”:

$$\begin{aligned} L_{\mathbf{G}}: \text{TM} &\rightarrow \mathbb{R} \\ v_x &\mapsto \frac{1}{2}\mathbf{G}(v_x, v_x). \end{aligned}$$

The following theorem is, of course, well-known, but we provide a proof for the purposes of making the paper self-contained.

8.7 Theorem: (Geodesics as extremals) *Let (M, \mathbf{G}) be a Riemannian manifold. For a C^2 -arc γ , the following are equivalent:*

- (i) γ is an $L_{\mathbf{G}}$ -extremal;
- (ii) γ is a geodesic of the Levi-Civita connection $\overset{\text{G}}{\nabla}$.

Proof: Suppose that $\gamma: [t_0, t_1] \rightarrow M$ is a C^2 -arc and let $\sigma: [-r, r] \times [t_0, t_1] \rightarrow M$ be a fixed endpoint variation of γ . We define

$$S_\sigma, T_\sigma: [-r, r] \times [t_0, t_1] \rightarrow \text{TM}$$

by

$$S_\sigma(s, t) = \frac{d}{ds}\sigma(s, t), \quad T_\sigma(s, t) = \frac{d}{dt}\sigma(s, t).$$

Thus $t \mapsto S_\sigma(s, t)$ should be thought of as a vector field along the curve σ_s , and $t \mapsto T_\sigma(s, t)$ should be thought of as the tangent vector field of the curve σ_s . The following lemma records a useful property of these two tangent vector fields along the curve σ_s .

1 Lemma: Let X_{S_σ} and X_{T_σ} be vector fields having the property that $X_{S_\sigma} \circ \sigma = S_\sigma$ and $X_{T_\sigma} \circ \sigma = T_\sigma$. Then $[X_{S_\sigma}, X_{T_\sigma}](\sigma(s, t)) = 0$ for $(s, t) \in [-r, r] \times [t_0, t_1]$.

Proof: Let f be a smooth function on M so that $f \circ \sigma: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{R}$. The definition of the Lie bracket gives

$$\begin{aligned} \mathcal{L}_{[X_{S_\sigma}, X_{T_\sigma}]} f \circ \sigma(s, t) &= \mathcal{L}_{X_{S_\sigma}} \mathcal{L}_{X_{T_\sigma}} f \circ \sigma(s, t) - \mathcal{L}_{X_{T_\sigma}} \mathcal{L}_{X_{S_\sigma}} f \circ \sigma(s, t) \\ &= \frac{d^2(f \circ \sigma)}{ds dt} - \frac{d^2(f \circ \sigma)}{dt ds}. \end{aligned}$$

Using the fact that mixed partial derivatives agree for C^2 -functions, the result now follows. \blacktriangledown

Note that, since $\overset{G}{\nabla}$ is torsion-free, the lemma implies that $\overset{G}{\nabla}_{S_\sigma} T_\sigma = \overset{G}{\nabla}_{T_\sigma} S_\sigma$. (This really means that $\overset{G}{\nabla}_{S_\sigma} T_\sigma(\sigma(s, t)) = \overset{G}{\nabla}_{T_\sigma} S_\sigma(\sigma(s, t))$ for all $(s, t) \in [-r, r] \times [t_0, t_1]$, just as in the lemma.) We also note that, since $\overset{G}{\nabla}$ is a metric connection, we have

$$\begin{aligned} \frac{d}{dt} \mathbf{G}(S_\sigma, T_\sigma)(s, t) &= \mathcal{L}_{T_\sigma} \mathbf{G}(S_\sigma, T_\sigma)(s, t) \\ &= \mathbf{G}(\overset{G}{\nabla}_{T_\sigma} S_\sigma, T_\sigma)(s, t) + \mathbf{G}(S_\sigma, \overset{G}{\nabla}_{T_\sigma} T_\sigma)(s, t) \end{aligned} \quad (8.2)$$

and

$$\frac{d}{ds} \mathbf{G}(T_\sigma, T_\sigma)(s, t) = \mathcal{L}_{S_\sigma} \mathbf{G}(T_\sigma, T_\sigma)(s, t) = 2\mathbf{G}(\overset{G}{\nabla}_{S_\sigma} T_\sigma, T_\sigma)(s, t). \quad (8.3)$$

Now we compute

$$\begin{aligned} \left. \frac{d}{ds} \right|_{s=0} \frac{1}{2} \int_{t_0}^{t_1} \mathbf{G}(\sigma'_s(t), \sigma'_s(t)) dt &= \left. \frac{d}{ds} \right|_{s=0} \frac{1}{2} \int_{t_0}^{t_1} \mathbf{G}(T_\sigma(s, t), T_\sigma(s, t)) dt \\ &= \int_{t_0}^{t_1} \mathbf{G}(\overset{G}{\nabla}_{S_\sigma(s, t)} T_\sigma(s, t), T_\sigma(s, t)) dt \Big|_{s=0} \\ &= \int_{t_0}^{t_1} \mathbf{G}(\overset{G}{\nabla}_{T_\sigma(s, t)} S_\sigma(s, t), T_\sigma(s, t)) dt \Big|_{s=0} \\ &= \int_{t_0}^{t_1} \mathbf{G}(\overset{G}{\nabla}_{\gamma'(t)} \delta\sigma(t), \gamma'(t)) dt \\ &= \int_{t_0}^{t_1} \left(\frac{d}{dt} \mathbf{G}(\delta\sigma(t), \gamma'(t)) - \mathbf{G}(\delta\sigma(t), \overset{G}{\nabla}_{\gamma'(t)} \gamma'(t)) \right) dt \\ &= \mathbf{G}(\delta\sigma(t), \gamma'(t)) \Big|_{t=t_0}^{t=t_1} - \int_{t_0}^{t_1} \mathbf{G}(\delta\sigma(t), \overset{G}{\nabla}_{\gamma'(t)} \gamma'(t)) dt \\ &= - \int_{t_0}^{t_1} \mathbf{G}(\delta\sigma(t), \overset{G}{\nabla}_{\gamma'(t)} \gamma'(t)) dt. \end{aligned}$$

Here we have used (8.3) in the third step, the lemma in the fourth step, (8.2) in the sixth step, and the vanishing of $\delta\sigma$ at the endpoints in the last step. Now, the expression

$$\int_{t_0}^{t_1} \mathbf{G}(\delta\sigma(t), \overset{G}{\nabla}_{\gamma'(t)} \gamma'(t)) dt$$

vanishes for every fixed endpoint variation σ if and only if $\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = 0$, by virtue of Lemma 8.3. \blacksquare

Next we work with an abstract setting for mechanics on a Riemannian manifold (M, \mathbb{G}) , which we think of as playing the rôle of $(Q_{0,\text{free}}, \mathbb{G}_{0,\mathcal{B}})$ for an interconnected rigid body system. As well as this manifold, we assume the following data:

1. a manifold M_0 with an injective immersion $\kappa: M_0 \rightarrow M$ (playing the rôle of Q_0);
2. the Riemannian metric $\mathbb{G}_0 = \kappa^*\mathbb{G}$ on M_0 (playing the rôle of \mathbb{G}_0);
3. the orthogonal projection $P_0: TM|M_0 \rightarrow TM_0$.
4. a time axis \mathbb{T} and a time interval $\mathbb{T}' \subseteq \mathbb{T}$;
5. a smooth mapping $F: \mathbb{T}' \times TM_0 \rightarrow T^*M$ such that $F(t, v_x) \in T_{\kappa(x)}^*M$ (playing the rôle of an external force);
6. the smooth mapping

$$\begin{aligned} F_0: \mathbb{T}' \times TM_0 &\rightarrow T^*M_0 \\ (x, v_x) &\mapsto P_0^* \circ F(t, v_x), \end{aligned}$$

where $P_0^*: T^*M|M_0 \rightarrow T^*M_0$ is the orthogonal projection defined by the metric \mathbb{G}^{-1} on the fibres of T^*M ;

7. a smooth distribution D on M_0 of locally constant rank (playing the rôle of the constraint distribution associated with velocity constraints).

The meaning of “playing the rôle of” will be made precise when we describe the Euler–Lagrange equations for an interconnected rigid body system. For now, we will just work with the above entities as abstractions of mechanical objects. We shall call the data $(M, M_0, \mathbb{G}, \mathbb{T}', F, D)$ an **abstract mechanical system**. Note that all other data above is determinable from this data. The associated **restricted abstract mechanical system** is $(M_0, \mathbb{G}_0, \mathbb{T}', F_0, D)$.

Given the preceding abstract data, we make the following definitions.

8.8 Definition: ((Restricted) forced and constrained geodesic equations) Given an abstract mechanical system $(M, M_0, \mathbb{G}, \mathbb{T}', F, D)$, a C^2 -curve $\gamma: \mathbb{T}' \rightarrow M_0$

- (i) satisfies the **forced and constrained geodesic equations** if $\gamma'(t) \in D_{\gamma(t)}$ for each $t \in \mathbb{T}'$ and if there exist smooth mappings

$$\iota, \lambda: \mathbb{T}' \rightarrow TM$$

satisfying the following conditions:

- (a) $\iota(t) \in (\text{image}(T_{\gamma(t)}\kappa))^\perp_{\mathbb{G}}$ for each $t \in \mathbb{T}'$;
- (b) $\lambda(t) \in (T_{\gamma(t)}\kappa(D_{\gamma(t)}))^\perp_{\mathbb{G}}$ for each $t \in \mathbb{T}'$;
- (c) the equation

$$\overset{\mathbb{G}}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma'(t) = \mathbb{G}^\sharp \circ F(t, \gamma'(t)) + \iota(t) + \lambda(t)$$

is satisfied for $t \in \mathbb{T}'$,

and

(ii) satisfies the **restricted forced and constrained geodesic equations** if $\gamma'(t) \in D_{\gamma(t)}$ for each $t \in \mathbb{T}'$ and if there exists a smooth mapping $\lambda_0: \mathbb{T}' \rightarrow TM_0$ satisfying the following conditions:

- (a) $\lambda_0(t) \in D_{\gamma(t)}^{\perp G}$ for each $t \in \mathbb{T}'$;
- (b) the equation

$$\overset{G_0}{\nabla}_{\gamma'(t)} \gamma'(t) = \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t)) + \lambda_0(t)$$

is satisfied for $t \in \mathbb{T}'$. •

Next we explore two alternative characterisations of the forced and constrained geodesic equations.

8.9 Proposition: (Lagrange–d’Alembert Principle (abstract versions)) *Let $(M, M_0, G, [t_0, t_1], F, D)$ be an abstract mechanical system. For a C^2 -arc $\gamma: [t_0, t_1] \rightarrow M_0$ satisfying $\gamma'(t) \in D_{\gamma(t)}$ for each $t \in [t_0, t_1]$, the following statements are equivalent:*

- (i) γ satisfies the forced and constrained geodesic equations;
- (ii) γ satisfies the **abstract instantaneous Lagrange–d’Alembert Principle**: for every $t \in [t_0, t_1]$,

$$\mathbf{G}(\overset{G}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma'(t)) - \mathbf{G}^\sharp \circ F(t, \gamma'(t)), v_x = 0, \quad v_x \in T_{\gamma(t)} \kappa(D_{\gamma(t)});$$

- (iii) γ satisfies the **abstract Lagrange–d’Alembert Principle**: for every fixed endpoint variation $\sigma: [-r, r] \times [t_0, t_1] \rightarrow M$ of $\kappa \circ \gamma$ satisfying $\delta\sigma(t) \in T_{\gamma(t)} \kappa(D_{\gamma(t)})$ for every $t \in [t_0, t_1]$,

$$\delta_\sigma L_G + \int_{t_0}^{t_1} \langle F(t, \gamma'(t)); \delta\sigma(t) \rangle dt = 0.$$

Proof: (i) \implies (ii) If γ satisfies the forced and constrained geodesic equations, let ι and λ satisfying

$$\iota(t) \in (\text{image}(T_{\gamma(t)} \kappa))^{\perp G}, \quad \lambda(t) \in (T_{\gamma(t)} \kappa(D_{\gamma(t)}))^{\perp G}, \quad t \in [t_0, t_1]$$

be the corresponding mappings. Since the forced and constrained geodesic equations hold

$$\overset{G}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma'(t) - \mathbf{G}^\sharp \circ F(t, \gamma'(t)) = \iota(t) + \lambda(t) \in (T_{\gamma(t)} \kappa(D_{\gamma(t)}))^{\perp G},$$

from which it immediately follows that (ii) holds.

(ii) \implies (i) Suppose that γ satisfies the abstract instantaneous Lagrange–d’Alembert Principle. For $t \in [t_0, t_1]$ we can write

$$T_{\kappa \circ \gamma(t)} M = \text{image}(T_{\gamma(t)} \kappa) \oplus (\text{image}(T_{\gamma(t)} \kappa))^{\perp G}.$$

We also have

$$\text{image}(T_{\gamma(t)} \kappa) = T_{\gamma(t)} \kappa(D_{\gamma(t)}) \oplus T_{\gamma(t)} \kappa(D_{\gamma(t)}^{\perp G_0}).$$

We then define vector fields ι and λ along $\kappa \circ \gamma$ by requiring $\iota(t)$ and $\lambda(t)$ to be the projections of $\overset{G}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma'(t) - \mathbf{G}^\sharp \circ F(t, \gamma'(t))$ onto $(\text{image}(T_{\gamma(t)} \kappa))^{\perp G}$ and $T_{\gamma(t)} \kappa(D_{\gamma(t)}^{\perp G_0})$, respectively.

Then, since $\overset{\mathbb{G}}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma(t) - \mathbf{G}^\sharp \circ X(t, \gamma'(t))$ is given as being \mathbf{G} -orthogonal to $T_{\gamma(t)} \kappa(D_{\gamma(t)})$, it follows that

$$\overset{\mathbb{G}}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma(t) - \mathbf{G}^\sharp \circ F(t, \gamma'(t)) = \iota(t) + \lambda(t),$$

as desired.

(i) \iff (iii) By the calculations from the proof of Theorem 8.7, we have

$$\delta_\sigma L_{\mathbf{G}} = - \int_{t_0}^{t_1} \mathbf{G}(\overset{\mathbb{G}}{\nabla}_{\kappa \circ \gamma'(t)} \kappa \circ \gamma(t), \delta\sigma(t)) dt.$$

The conclusion now follows directly from the definition of the forced and constrained geodesic equations. \blacksquare

We can also formulate the conditions of the preceding result directly on M_0 .

8.10 Proposition: (Lagrange–d’Alembert Principle (restricted abstract versions))

Let $(M, M_0, \mathbf{G}, [t_0, t_1], F, D)$ be an abstract mechanical system with $(M_0, \mathbf{G}_0, [t_0, t_1], F_0, D)$ the corresponding restricted abstract mechanical system. For a C^2 -arc $\gamma: [t_0, t_1] \rightarrow M_0$ satisfying $\gamma'(t) \in D_{\gamma(t)}$ for each $t \in [t_0, t_1]$, the following statements are equivalent:

- (i) γ satisfies the forced and constrained geodesic equations;
- (ii) γ satisfies the restricted forced and constrained geodesic equations;
- (iii) γ satisfies the **restricted abstract instantaneous Lagrange–d’Alembert Principle**: for every $t \in [t_0, t_1]$,

$$\mathbf{G}_0(\overset{\mathbb{G}_0}{\nabla}_{\gamma'(t)} \gamma'(t) - \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t)), v_x) = 0, \quad v_x \in D_{\gamma(t)};$$

- (iv) γ satisfies the **restricted abstract Lagrange–d’Alembert Principle**: for every fixed endpoint variation $\sigma_0: [-r, r] \times [t_0, t_1] \rightarrow M_0$ of γ satisfying $\delta\sigma_0(t) \in D_{\gamma(t)}$ for every $t \in [t_0, t_1]$,

$$\delta_\sigma L_{\mathbf{G}_0} + \int_{t_0}^{t_1} \langle F_0(t, \gamma'(t)); \delta\sigma_0(t) \rangle dt = 0.$$

Proof: The equivalence of the last three statements follows from Proposition 8.9 with $M_0 = M$. To complete the proof we will show the equivalence of parts (i) and (ii). Since both of these statements are locally determined, we may suppose that $\kappa(M_0)$ is an embedded (but not necessarily properly embedded) submanifold of M . We thus identify M_0 with $\kappa(M_0)$. We also suppose that, in some neighbourhood of M_0 , M_0 is an integral manifold of some integrable distribution \mathbb{I} . We are thus in a position to use Lemma 8.6, and we note that, for vector fields X and Y on M_0 that we locally extend to vector fields on M , we have

$$\overset{\mathbb{G}_0}{\nabla}_X Y = \overset{\mathbb{G}}{\nabla}_X Y + (\overset{\mathbb{G}}{\nabla}_X P_1^\perp)(Y),$$

where P_1^\perp is the \mathbf{G} -orthogonal projection onto \mathbb{I}^\perp .

(i) \implies (ii) Suppose that γ satisfies the forced and constrained geodesic equations and let $\iota: [t_0, t_1] \rightarrow \mathbf{T}M$ and $\lambda: [t_0, t_1] \rightarrow \mathbf{T}M$ be the corresponding mappings. Let us write $\lambda(t) = \lambda_0(t) + \lambda_1(t)$ where

$$\lambda_0(t) \in T_{\gamma(t)} M_0, \quad \lambda_1(t) \in (T_{\gamma(t)} M_0)^\perp.$$

We thus have

$$\overset{\mathbb{G}_0}{\nabla}_{\gamma'(t)}\gamma'(t) - (\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_1^\perp)(\gamma'(t)) = \mathbb{G}^\sharp \circ F(t, \gamma'(t)) + \iota(t) + \lambda(t).$$

Taking the \mathbb{G} -orthogonal component of this equation tangent to TM_0 gives

$$\overset{\mathbb{G}_0}{\nabla}_{\gamma'(t)}\gamma'(t) = P_1 \circ \mathbb{G}^\sharp \circ F(t, \gamma'(t)) + \lambda_0(t),$$

where P_1 is the \mathbb{G} -orthogonal projection onto \mathbb{I} . Thus this part of the result follows after we note that

$$P_1 \circ \mathbb{G}^\sharp \circ F(t, \gamma'(t)) = \mathbb{G}_0^\sharp \circ F_0(t, \gamma'(t)). \quad (8.4)$$

(ii) \implies (i) Let γ satisfy the restricted forced and constrained geodesic equations, and let $\lambda: [t_0, t_1] \rightarrow \text{TM}_0$ be the corresponding mapping. We thus have

$$\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}\gamma'(t) = -(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_1^\perp)(\gamma'(t)) + \mathbb{G}_0 \circ F_0(t, \gamma'(t)) + \lambda(t).$$

Keeping in mind (8.4), this part of the result follows by taking $\iota(t) = -(\overset{\mathbb{G}}{\nabla}_{\gamma'(t)}P_1^\perp)(\gamma'(t))$. ■

We have been rather pedantic in the preceding development about the rôle of the inclusion κ , carefully *not* identifying M_0 with its image under κ . This is really necessary since $\kappa(\text{M}_0)$ is only an immersed submanifold.

This turns out to be a convenient setting to discuss the existence and uniqueness of solutions to forced and constrained mechanics. Let us do this so as to resolve this issue in the cases where it has been resolved in the literature. In doing this, we see that we *must* assume that D is smooth and locally constant rank. We have the following result, essentially following from [Lewis 1998, Theorem 5.4].

8.11 Theorem: (An affine connection characterisation of the forced and constrained geodesic equations) *Let $(\text{M}, \text{M}_0, \mathbb{G}, \mathbb{T}', F, \text{D})$ be an abstract mechanical system. For a C^2 -curve $\gamma: \mathbb{T}' \rightarrow \text{M}_0$, the following statements are equivalent:*

- (i) γ satisfies the forced and constrained geodesic equations;
- (ii) $\gamma'(t_0) \in \text{D}_{\gamma(t_0)}$ for some $t_0 \in \mathbb{T}'$ and γ satisfies the equation

$$\overset{\text{D}}{\nabla}_{\gamma'(t)}\gamma'(t) = P_{\text{D}} \circ \mathbb{G}_0^\sharp \circ F_0(t, \gamma'(t)), \quad t \in \mathbb{T}'. \quad (8.5)$$

In particular, the forced and constrained geodesic equations possess uniquely determined solutions locally in time.

Proof: By Proposition 8.10 we shall work with the restricted forced and constrained geodesic equations.

Thus we first suppose that γ satisfies the forced and constrained geodesic equations, so there exists $\lambda: \mathbb{T}' \rightarrow \text{D}^{\perp \mathbb{G}_0}$ such that

$$\overset{\text{D}}{\nabla}_{\gamma'(t)}\gamma'(t) = (\overset{\mathbb{G}_0}{\nabla}_{\gamma'(t)}P_{\text{D}}^\perp)(\gamma'(t)) + \mathbb{G}_0^\sharp \circ F_0(t, \gamma'(t)) + \lambda(t).$$

Taking the D -component of the \mathbb{G}_0 -orthogonal decomposition, and using Lemma 8.5(i), gives

$$\overset{\text{D}}{\nabla}_{\gamma'(t)}\gamma'(t) = P_{\text{D}} \circ \mathbb{G}_0^\sharp \circ F_0(t, \gamma'(t)).$$

Moreover, since $\gamma'(t) \in D_{\gamma(t)}$ for all $t \in \mathbb{T}'$ by hypothesis, it certainly holds that $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some $t_0 \in \mathbb{T}'$.

Now suppose that $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some $t_0 \in \mathbb{T}'$ and that

$$\overset{\text{D}}{\nabla}_{\gamma'(t)} \gamma'(t) = P_D \circ \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t)).$$

Then

$$\overset{\text{G}}{\nabla}_{\gamma'(t)} \gamma'(t) = -(\overset{\text{G}}{\nabla}_{\gamma'(t)} P_D^\perp)(\gamma'(t)) + \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t)) - P_D^\perp \circ \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t)),$$

using $P_D = \text{id}_{\text{TM}_0} - P_D^\perp$. Taking

$$\lambda(t) = -(\overset{\text{G}}{\nabla}_{\gamma'(t)} P_D^\perp)(\gamma'(t)) - P_D^\perp \circ \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t))$$

gives

$$\overset{\text{G}}{\nabla}_{\gamma'(t)} \gamma'(t) = \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t)) + \lambda(t),$$

and $\lambda(t) \in D^\perp_{G_0}$ by Lemma 8.5(ii).

It remains to show that $\gamma'(t) \in D_{\gamma(t)}$ for all $t \in \mathbb{T}'$. That is to say, thought of as a curve in TM_0 , we show that $\gamma'(t) \in D \subseteq \text{TM}_0$ if $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some $t_0 \in \mathbb{T}'$. To do this, let Z denote the geodesic spray associated with the affine connection $\overset{\text{D}}{\nabla}$. Thus Z is the second-order vector field on TM_0 for which the projections of integral curves to M_0 are geodesics of $\overset{\text{D}}{\nabla}$. We also denote

$$\text{vlft}(P_D \circ \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t))) = \left. \frac{d}{ds} \right|_{s=0} (\gamma'(t) + s P_D \circ \mathbf{G}_0^\sharp \circ F_0(t, \gamma'(t))) \in \mathcal{T}_{\gamma'(t)} \text{TM}.$$

We then note that we can write (8.5) in the form

$$\Upsilon'(t) = Z(\Upsilon(t)) - \text{vlft}(P_D \circ \mathbf{G}_0^\sharp \circ F_0(t, \Upsilon(t))) \quad (8.6)$$

for a curve $\Upsilon: \mathbb{T}' \rightarrow \text{TM}_0$ [Bullo and Lewis 2004, Equation (4.25)]. We abbreviate $\gamma = \pi_{\text{TM}} \circ \Upsilon$. First note that, since

$$P_D \circ \mathbf{G}_0^\sharp \circ F_0(t, \Upsilon(t)) \in D_{\gamma(t)}, \quad t \in \mathbb{T}',$$

it follows from the definition that $\text{vlft}(P_D \circ \mathbf{G}_0^\sharp \circ F_0(t, \Upsilon(t)))$ is tangent to D . By [Lewis 1998, Theorem 5.4] the tangency of Z to D will follow if we can show that $\overset{\text{D}}{\nabla}_Y Y$ is D -valued for every D -valued vector field Y . This, however, follows from Lemma 8.5(i). We thus conclude that any solution to (8.6) for which $\Upsilon(t_0) \in D$ for some $t_0 \in \mathbb{T}'$ will have the property that $\Upsilon(t) \in D$ for all $t \in \mathbb{T}'$. Therefore, solutions of (8.5) with $\gamma'(t_0) \in D_{\gamma(t_0)}$ for some $t_0 \in \mathbb{T}'$ will satisfy $\gamma'(t) \in D_{\gamma(t)}$ for all $t \in \mathbb{T}'$.

The final assertion of the theorem follows from the fact that the equation (8.6), whose solutions with initial conditions in D are exactly the solutions of the restricted forced and constrained geodesic equations, is just that for integral curves of a smooth time-varying vector field. \blacksquare

As mentioned in the introduction, when D is not smooth or of locally constant rank, the matter of existence and uniqueness of solutions to the forced and constrained geodesic equations that arise from mechanics—see Theorem 8.18 below—is a subject that requires further research.

8.4. Euler–Lagrange equations for a single, free, unforced rigid body. A key pivot point for our transition from Newton–Euler to Euler–Lagrange is the means by which this is done in the simplest possible case, that of a single rigid body, with no constraints or forces imposed upon it.

First let us characterise extremals for a particular Lagrangian on $\mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S}))$. We define

$$\begin{aligned} L_{\mathcal{B}}: \mathbb{T}(\text{Rgd}(\mathcal{B}; \mathcal{S})) &\rightarrow \mathbb{R} \\ (\Phi, (A, v)) &\mapsto \frac{1}{2}G_{\mathcal{B}}((A, v), (A, v)), \end{aligned}$$

i.e., the kinetic energy for the body. We then have the following result.

8.12 Lemma: (Solutions of the Newton–Euler equations as extremals) *Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathfrak{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . For a C^2 -arc $\phi: [t_0, t_1] \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$, the following statements are equivalent:*

- (i) ϕ is an $L_{\mathcal{B}}$ -extremal;
- (ii) ϕ satisfies the Newton–Euler equations.

Proof: We shall choose a spatial origin x_0 so that, when convenient, we can represent $\Phi \in \text{Rgd}(\mathcal{B}; \mathcal{S})$ with $(R, r) \in \text{Isom}^+(\mathcal{B}; \mathcal{S}) \times \mathbb{V}$. Our arguments, however, do not in the end depend on this choice.

Define

$$\begin{aligned} \ell_{\mathcal{B}}: \mathbb{U} \oplus \mathbb{U} &\rightarrow \mathbb{R} \\ (\Omega, V) &\mapsto \frac{1}{2}G(\mathbb{I}_c(\Omega), \Omega) + \frac{1}{2}mG(V, V). \end{aligned}$$

Note that

$$L_{\mathcal{B}}(\Phi, (A, v)) = \ell_{\mathcal{B}}(\widetilde{R_{\Phi}^T \circ A}, R_{\Phi}^T(v)).$$

Let $\phi: [t_0, t_1] \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ be a C^2 -arc and let σ be a fixed endpoint variation of ϕ . Let us adapt the notation of Section 6.5 and define

$$\Omega_{\sigma}(s, t) = R_{\sigma}^T(s, t) \circ \widetilde{\dot{R}_{\sigma}(s, t)}, \quad V_{\sigma}(s, t) = R_{\sigma}^T(s, t)(v_{\sigma}(s, t)).$$

Let us also define

$$\begin{aligned} \delta R_{\sigma}(t) &= \left. \frac{d}{ds} \right|_{s=0} R_{\sigma}(s, t), \quad \delta r_{\sigma}(t) = \left. \frac{d}{ds} \right|_{s=0} r_{\sigma}(s, t), \quad \delta \Omega_{\sigma}(t) = \left. \frac{d}{ds} \right|_{s=0} \Omega_{\sigma}(s, t), \\ \delta v_{\sigma}(t) &= \left. \frac{d}{ds} \right|_{s=0} v_{\sigma}(s, t), \quad \delta V_{\sigma}(t) = \left. \frac{d}{ds} \right|_{s=0} V_{\sigma}(s, t). \end{aligned}$$

Time-derivatives we denote with “dots”:

$$\begin{aligned} \dot{R}_{\sigma}(s, t) &= \frac{d}{dt} R_{\sigma}(s, t), \quad \dot{r}_{\sigma}(s, t) = \frac{d}{dt} r_{\sigma}(s, t) = v_{\sigma}(s, t), \quad \dot{\Omega}_{\sigma}(s, t) = \frac{d}{dt} \Omega_{\sigma}(s, t), \\ \dot{v}_{\sigma}(s, t) &= \frac{d}{dt} v_{\sigma}(s, t), \quad \dot{V}_{\sigma}(s, t) = \frac{d}{dt} V_{\sigma}(s, t). \end{aligned}$$

We will adopt the notational convention that evaluation at $(s, t) = (0, t)$ will simply be abbreviated with evaluation at t . We also freely interchange s - and t -derivatives since all functions are of class C^2 and take values in finite-dimensional vector spaces.

Let us make some preliminary computations. We define

$$\Xi_{\sigma}(t) = R_{\sigma}^T(t) \circ \widetilde{\delta R_{\sigma}(t)}, \quad X_{\sigma}(t) = R_{\sigma}^T(t)(\delta r_{\sigma}(t)).$$

Next we calculate

$$\frac{d}{dt}\widehat{\Xi}_\sigma(t) = \dot{R}_\sigma^T(t) \circ \delta R_\sigma(t) + R_\sigma^T(t) \circ \delta \dot{R}_\sigma(t)$$

and

$$\frac{d}{dt}X_\sigma(t) = \dot{R}_\sigma^T(t)(\delta r_\sigma(t)) + R_\sigma^T(t)(\delta v_\sigma(t)).$$

We calculate

$$\delta \widehat{\Omega}_\sigma(t) = \delta R_\sigma^T(t) \circ \dot{R}_\sigma(t) + R_\sigma^T(t) \circ \delta \dot{R}_\sigma(t)$$

and

$$\delta V_\sigma(t) = \delta R_\sigma^T(t)(v_\sigma(t)) + R_\sigma^T(t)(\delta v_\sigma(t)).$$

Combining these two expressions with the two preceding them, we get

$$\delta \widehat{\Omega}_\sigma(t) = \frac{d}{dt}\widehat{\Xi}_\sigma(t) + \delta R_\sigma^T(t) \circ \dot{R}_\sigma(t) - \dot{R}_\sigma^T(t) \circ \delta R_\sigma(t)$$

and

$$\delta V_\sigma(t) = \frac{d}{dt}X_\sigma(t) + \delta R_\sigma^T(t)(v_\sigma(t)) - \dot{R}_\sigma^T(t)(\delta r_\sigma(t)).$$

The definitions of the symbols involved gives

$$\delta R_\sigma^T(t) \circ \dot{R}_\sigma(t) - \dot{R}_\sigma^T(t) \circ \delta R_\sigma(t) = \widehat{\Omega}_\sigma(t) \circ \widehat{\Xi}_\sigma(t) - \widehat{\Xi}_\sigma(t) \circ \widehat{\Omega}_\sigma(t)$$

and

$$\delta R_\sigma^T(t)(v_\sigma(t)) - \dot{R}_\sigma^T(t)(\delta r_\sigma(t)) = \Omega_\sigma(t) \times X_\sigma(t) - \Xi_\sigma(t) \times V_\sigma(t).$$

Thus we have

$$\delta \Omega_\sigma(t) = \dot{\Xi}_\sigma(t) + \Omega_\sigma(t) \times \Xi_\sigma(t) \tag{8.7}$$

and

$$\delta V_\sigma(t) = \dot{X}_\sigma(t) + \Omega_\sigma(t) \times X_\sigma(t) - \Xi_\sigma(t) \times V_\sigma(t), \tag{8.8}$$

using the identity

$$\widehat{U} \circ \widehat{V} - \widehat{V} \circ \widehat{U} = \widehat{U \times V}, \quad U, V \in \mathbf{U},$$

for the first equation, which can be proved using the Jacobi identity for the cross-product.

With these preliminary computations out of the way, we may now proceed with the

main part of the proof. We calculate

$$\begin{aligned}
\delta L_{\mathcal{B}} &= \left. \frac{d}{ds} \right|_{s=0} \int_{t_0}^{t_1} L_{\mathcal{B}}(\sigma(s, t), (\dot{R}_{\sigma}(s, t), v_{\sigma}(s, t))) dt \\
&= \int_{t_0}^{t_1} \left. \frac{d}{ds} \right|_{s=0} \ell_{\mathcal{B}}(\Omega_{\sigma}(s, t), V_{\sigma}(s, t)) dt \\
&= \int_{t_0}^{t_1} (G(\mathbb{I}_c(\Omega_{\sigma}(t)), \delta\Omega_{\sigma}(t)) + mG(V_{\sigma}(t), \delta V_{\sigma}(t))) dt \\
&= \int_{t_0}^{t_1} (G(\mathbb{I}_c(\Omega_{\sigma}(t)), \dot{\Xi}_{\sigma}(t) + \Omega_{\sigma}(t) \times \Xi_{\sigma}(t)) \\
&\quad + mG(V_{\sigma}(t), \dot{X}_{\sigma}(t) + \Omega_{\sigma}(t) \times X_{\sigma}(t) - \Xi_{\sigma}(t) \times V_{\sigma}(t))) dt \\
&= \int_{t_0}^{t_1} (G(\mathbb{I}_c(\Omega_{\sigma}(t)), \dot{\Xi}_{\sigma}(t)) + G((\mathbb{I}_c(\Omega_{\sigma}(t))) \times \Omega_{\sigma}(t), \Xi_{\sigma}(t)) \\
&\quad + mG(V_{\sigma}(t), \dot{X}_{\sigma}(t)) + mG(V_{\sigma}(t) \times \Omega_{\sigma}(t), X_{\sigma}(t))) dt \\
&= G(\mathbb{I}_c(\Omega_{\sigma}(t)), \Xi_{\sigma}(t)) \Big|_{t=t_0}^{t=t_1} + mG(V_{\sigma}(t), X_{\sigma}(t)) \Big|_{t=t_0}^{t=t_1} \\
&\quad - \int_{t_0}^{t_1} (G(\frac{d}{dt}(\mathbb{I}_c(\Omega_{\sigma}(t))) - \mathbb{I}_c(\Omega_{\sigma}(t)) \times \Omega_{\sigma}(t), \Xi_{\sigma}(t)) \\
&\quad + G(\frac{d}{dt}(mV_{\sigma}(t)) - V_{\sigma}(t) \times \Omega_{\sigma}(t), X_{\sigma}(t))) dt \\
&= - \int_{t_0}^{t_1} (G(\frac{d}{dt}(\mathbb{I}_c(\Omega_{\sigma}(t))) + \Omega_{\sigma}(t) \times (\mathbb{I}_c(\Omega_{\sigma}(t))), \Xi_{\sigma}(t)) \\
&\quad + G(\frac{d}{dt}(mV_{\sigma}(t)) + \Omega_{\sigma}(t) \times V_{\sigma}(t), X_{\sigma}(t))) dt,
\end{aligned}$$

using the formulae (8.7) and (8.8), equation (1.1), integration by parts, and the fact that Ξ_{σ} and X_{σ} vanish at the endpoints of the interval $[t_0, t_1]$. An easy argument, using Lemma 8.3, shows that, for any $\Xi, X: [t_0, t_1] \rightarrow \mathbb{U}$ that vanish at the endpoints, there exists a fixed endpoint variation σ of ϕ such that $\Xi = \Xi_{\sigma}$ and $X = X_{\sigma}$. This proves the lemma by virtue of Theorem 7.4. \blacksquare

Now we may provide an important step in our path from Newton–Euler to Euler–Lagrange.

8.13 Theorem: (Solutions of the unforced Newton–Euler equations as geodesics (single body version)) *Let $\mathcal{S} = (\mathbb{S}, \mathbb{V}, \mathbb{g}, \theta)$ be a Newtonian space model, let $\mathcal{B} = (\mathbb{B}, \mathbb{U}, \mathbb{G}, \Theta)$ be a body reference space, and let (\mathcal{B}, μ) be a rigid body in \mathcal{B} . For a C^2 -rigid motion $\phi: \mathbb{T}' \rightarrow \text{Rgd}(\mathcal{B}; \mathcal{S})$ the following statements are equivalent:*

- (i) ϕ satisfies the unforced Newton–Euler equations for the body (\mathcal{B}, μ) ;
- (ii) $\phi_0 \triangleq \pi_{\mathcal{B}} \circ \phi$ is a geodesic for the Riemannian metric $\mathbb{G}_{0, \mathcal{B}}$.

Proof: Since the properties (i) and (ii) are both local, in the proof we can suppose that \mathbb{T}' is sufficiently small that ϕ is a C^2 -arc. We can thus assume, without loss of generality, that $\mathbb{T}' = [t_0, t_1]$. For brevity of notation, let us denote

$$\mathbb{Q} = \text{Rgd}(\mathcal{B}; \mathcal{S}), \quad \mathbb{Q}_0 = \text{Rgd}(\mathcal{B}; \mathcal{S})/\mathbb{I}(\mathcal{B}), \quad \mathbb{H} = \mathbb{I}(\mathcal{B}).$$

Apart from brevity, this notation will show that certain of our arguments are rather general regarding Lie group actions on manifolds.

(i) \implies (ii) We suppose that we have a smooth connection on the bundle $\pi_0: \mathbb{Q} \rightarrow \mathbb{Q}_0$ by [Kobayashi and Nomizu 1963, Theorem II.2.1]. Suppose that $\phi: [t_0, t_1] \rightarrow \mathbb{Q}$ satisfies the Newton–Euler equations. Let $\sigma_0: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{T}\mathbb{Q}_0$ be a fixed endpoint variation of ϕ_0 . Define $\bar{\sigma}: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{Q}$ by asking that $t \mapsto \bar{\sigma}(s, t)$ be the horizontal lift of $t \mapsto \sigma_0(s, t)$ through $\phi(t_0)$ for each $s \in [-r, r]$. Since horizontal lift is obtained by solving differential equations—cf. the proof of [Kobayashi and Nomizu 1963, Proposition II.3.1]—and since solutions of differential equations depend regularly on initial condition, one can easily show that $\bar{\sigma}$ is of class C^2 . Let $\psi_1: [t_0, t_1] \rightarrow \mathbb{H}$ be a C^2 -curve such that $\bar{\sigma}(0, t) \circ \psi_1(t) = \phi(t)$ and such that $\psi_1(t_0) = \text{id}$. This is possible since $\pi_0 \circ \bar{\sigma}(0, t) = \pi_0 \circ \phi(t)$ and by regularity of solutions of differential equations with respect to initial conditions. Let $\psi_2: [-r, r] \rightarrow \mathbb{H}$ be a C^2 -curve in \mathbb{H} such that $\bar{\sigma}(s, t_1) \circ \psi_1(t_1) \circ \psi_2(s) = \phi(t_1)$ and such that $\psi_2(0) = \text{id}$. Such a curve with the prescribed regularity exists, again by regularity of solutions to differential equations with respect to initial conditions. Now define

$$\begin{aligned} \sigma: [-r, r] \times [t_0, t_1] &\rightarrow \mathbb{Q} \\ (s, t) &\mapsto \bar{\sigma}(s, t) \circ \psi_1(t) \circ \psi_2\left(\frac{t-t_0}{t_1-t_0}s\right). \end{aligned}$$

We need to calculate derivatives of σ with respect to t and s . To make this computation more convenient, we shall use a spatial origin, along with our body origin at the centre of mass. We then write elements of \mathbb{Q} as matrices:

$$\Phi \sim \begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix}, \quad R \in \text{Isom}^+(\mathcal{B}; \mathcal{S}), \quad r \in \mathbb{V}.$$

We similarly write elements of $\text{Rgd}(\mathcal{B})$ as matrices:

$$\Psi \sim \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix}, \quad A \in \text{SO}(\mathbb{U}, \mathbb{G}, \Theta), \quad a \in \mathbb{U}.$$

The right-action of $\text{Rgd}(\mathcal{B})$ on $\text{Rgd}(\mathcal{B}; \mathcal{S})$ is then by matrix multiplication on the right:

$$\left(\begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} \right) \mapsto \begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix}.$$

We shall write

$$\bar{\sigma}(s, t) = \begin{bmatrix} \bar{R}(s, t) & \bar{r}(s, t) \\ 0 & 1 \end{bmatrix}.$$

Since \mathbb{H} is a subgroup of the group of rotations fixing X_c , we easily see that a vertical tangent vector at

$$\begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix}$$

will have the form

$$\begin{bmatrix} R & r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\Omega} & 0 \\ 0 & 0 \end{bmatrix}$$

for some $\Omega \in \mathbb{U}$. We write

$$\psi_1(t) = \begin{bmatrix} A_1(t) & 0 \\ 0 & 1 \end{bmatrix}, \quad \psi_2(s) = \begin{bmatrix} A_2(s) & 0 \\ 0 & 1 \end{bmatrix}$$

for curves $t \mapsto A_1(t)$ and $s \mapsto A_2(s)$ in $\text{SO}(\mathbf{U}; \mathbf{G}, \Theta)$. Thus

$$\begin{aligned} \frac{d}{dt}\sigma(s, t) &= \begin{bmatrix} \frac{d}{dt}\bar{R}(s, t) & \frac{d}{dt}\bar{r}(s, t) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1(t) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 1 \end{bmatrix} \\ &\quad + \begin{bmatrix} \bar{R}(s, t) & \bar{r}(s, t) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{A}_1(t) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 1 \end{bmatrix} \\ &\quad + \frac{s}{t_1-t_0} \begin{bmatrix} \bar{R}(s, t) & \bar{r}(s, t) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1(t) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{A}_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

If we define

$$\hat{\Omega}_1(t) = A_1^T(t) \circ \dot{A}_1(t), \quad \hat{\Omega}_2(t) = A_2^T(t) \circ \dot{A}_2(t),$$

then we can rewrite this as

$$\begin{aligned} \frac{d}{dt}\sigma(s, t) &= \underbrace{\begin{bmatrix} \frac{d}{dt}\bar{R}(s, t) & \frac{d}{dt}\bar{r}(s, t) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1(t) \circ A_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 1 \end{bmatrix}}_{\text{projects to } \sigma'_0(s, t)} \\ &\quad + \underbrace{\begin{bmatrix} \bar{R}(s, t) & \bar{r}(s, t) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1(t) \circ A_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_2^T(\frac{t-t_0}{t_1-t_0}s)(\hat{\Omega}_1(t)) & 0 \\ 0 & 0 \end{bmatrix}}_{\text{vertical}} \\ &\quad + \frac{s}{t_1-t_0} \underbrace{\begin{bmatrix} \bar{R}(s, t) & \bar{r}(s, t) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1(t) \circ A_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\Omega}_2(\frac{t-t_0}{t_1-t_0}s) & 0 \\ 0 & 0 \end{bmatrix}}_{\text{vertical}}, \quad (8.9) \end{aligned}$$

using Lemma 2.17. In like manner we compute

$$\begin{aligned} \frac{d}{ds}\sigma(0, t) &= \underbrace{\begin{bmatrix} \frac{d}{ds}\bar{R}(0, t) & \frac{d}{ds}\bar{r}(0, t) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1(t) & 0 \\ 0 & 1 \end{bmatrix}}_{\text{projects to } \delta\sigma_0(t)} \\ &\quad + \frac{t-t_0}{t_1-t_0} \underbrace{\begin{bmatrix} \bar{R}(s, t) & \bar{r}(s, t) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1(t) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\Omega}_2(0) & 0 \\ 0 & 0 \end{bmatrix}}_{\text{vertical}}. \quad (8.10) \end{aligned}$$

We can then directly verify the following properties of σ :

1. σ is of class C^2 ;
2. $\sigma(s, t_0) = \phi(t_0)$ and $\sigma(s, t_1) = \phi(t_1)$, $s \in [-r, r]$;
3. $\pi_0 \circ \sigma(s, t) = \sigma_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$;
4. $T_{\sigma(s, t)}\pi_0(\sigma'(s, t)) = \sigma'_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$.

We then have $L_{\mathcal{B}}(\sigma'(s, t)) = L_{\mathbf{G}_{0, \mathcal{B}}}(\sigma'_0(s, t))$ (by definition of $\mathbf{G}_{0, \mathcal{B}}$) which gives

$$\frac{d}{ds} \Big|_{s=0} \int_{t_0}^{t_1} L_{\mathcal{B}}(\sigma'(s, t)) dt = \frac{d}{ds} \Big|_{s=0} \int_{t_0}^{t_1} L_{\mathbf{G}_{0, \mathcal{B}}}(\sigma_0(s, t)) dt,$$

and so this part of the theorem follows from Theorem 8.7 and Lemma 8.12.

(ii) \implies (i) Suppose that ϕ_0 is a geodesic for $\mathbf{G}_{0, \mathcal{B}}$. Let $\sigma: [-r, r] \times [t_0, t_1] \rightarrow \mathbf{Q}$ be a fixed endpoint variation of ϕ and define $\sigma_0 = \pi_0 \circ \sigma$. Then one readily verifies that:

1. σ_0 is of class C^2 ;
2. $\sigma_0(s, t_0) = \phi_0(t_0)$ and $\sigma_0(s, t_1) = \phi_0(t_1)$, $s \in [-r, r]$;
3. $T_{\sigma(s,t)}\pi_0(\sigma'(s, t)) = \sigma'_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$.

Then, just as in the first part of the proof,

$$\frac{d}{ds} \Big|_{s=0} \int_{t_0}^{t_1} L_{\mathcal{B}}(\sigma'(s, t)) dt = \frac{d}{ds} \Big|_{s=0} \int_{t_0}^{t_1} L_{\mathbb{G}_0, \mathcal{B}}(\sigma_0(s, t)) dt,$$

and the theorem follows from Theorem 8.7 and Lemma 8.12. \blacksquare

8.5. Lagrangian representations of torque-forces. A crucial element in our development of the Euler–Lagrange equations from the Newton–Euler equations is the precise understanding of how forces and torques in the Newton–Euler setting, i.e., those from Definition 6.11, are translated to a geometrically meaningful object on the configuration manifold of the system.

The key is the following result.

8.14 Lemma: (Lagrangian representation of torque-force) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . If $\boldsymbol{\tau} \oplus \mathbf{f}$ is a primary torque-force on \mathcal{B} at $\Phi \in \mathcal{Q}$, then there exists a unique $F \in \mathbb{T}_{\Phi}^* \mathcal{Q}_{\text{free}}$ such that*

$$\langle F; \psi'(0) \rangle = \sum_{a=1}^m (\mathfrak{g}(\tau_a, \omega_{\psi_a}(0)) + \mathfrak{g}(f_a, v_{\psi_a}(0)))$$

for every differentiable motion $\psi: [-r, r] \rightarrow \mathcal{Q}_{\text{free}}$ of the free system. Moreover, $F \in \text{ann}(\ker(T_{\Phi}\pi_0))$ and so defines $F_0 \in \mathbb{T}_{\pi_0(\Phi)}^* \mathcal{Q}_{0, \text{free}}$ satisfying

$$\langle F_0; T_{\Phi}\pi_0(\mathbf{A}, \mathbf{v}) \rangle = \langle F; (\mathbf{A}, \mathbf{v}) \rangle \quad (8.11)$$

for $(\mathbf{A}, \mathbf{v}) \in \mathbb{T}_{\Phi} \mathcal{Q}_{\text{free}}$.

Proof: The existence and uniqueness of F follows since the maps

$$\mathbb{T}_{\Phi} \mathcal{Q}_{\text{free}} \ni (\mathbf{A}, \mathbf{v}) \mapsto ((A_1 \circ R_{\Phi_1}^T, \dots, A_m \circ R_{\Phi_m}^T), (v_1, \dots, v_m)) \in \mathbb{V}^m \oplus \mathbb{V}^m$$

and

$$\mathbb{V}^m \oplus \mathbb{V}^m \ni (\boldsymbol{\omega}, \mathbf{v}) \mapsto \sum_{a=1}^m (\mathfrak{g}(\tau_a, \omega_a) + \mathfrak{g}(f_a, v_a))$$

are linear, and thus so is their composition. Thus the right-hand side of the expression defining F in the statement of the lemma does indeed define a linear function on $\mathbb{T}_{\Phi} \mathcal{Q}_{\text{free}}$.

Since $\boldsymbol{\tau} \oplus \mathbf{f}$ is a primary torque-force, $\tau_a \in (\ker(R_{\Phi_a} \circ \mathbb{I}_{c,a} \circ R_{\Phi_a}^T))^{\perp \mathfrak{g}}$. The definition of \mathbb{G} and Lemma 4.12 then give $F \in \text{ann}(\ker(T_{\Phi}\pi_0))$. That this defines $F_0 \in \mathbb{T}_{\pi_0(\Phi)}^* \mathcal{Q}_{0, \text{free}}$ satisfying (8.11) follows from two observations.

1. By Proposition 4.15 we have $\mathbb{T}_{\pi_0(\Phi)} \mathcal{Q}_{0, \text{free}} \simeq \mathbb{T}_{\Phi} \mathcal{Q}_{\text{free}} / \ker(T_{\Phi}\pi_0)$.

2. We have the following general fact. If X is a \mathbb{R} -vector space and $Y \subseteq X$ is a subspace, then we have a natural isomorphism $\iota: \text{ann}(Y) \rightarrow (X/Y)^*$ defined by

$$\langle \iota(\alpha); x + Y \rangle = \langle \alpha; x \rangle, \quad \alpha \in \text{ann}(Y), \quad x \in X. \quad \blacksquare$$

A reference to Section 6.5 will lead the reader to note that our characterisation of the cotangent vector F is made by using the device of virtual power.

Let us give names to the objects constructed in the lemma.

8.15 Definition: (Lagrangian force) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold Q . Let $\tau \oplus \mathbf{f}$ be a primary torque-force on \mathcal{B} at $\Phi \in Q$.

- (i) The element $F \in T_{\Phi}^* Q_{\text{free}}$ characterised in Lemma 8.14 is the *Lagrangian force* associated to $\tau \oplus \mathbf{f}$.
- (ii) The element $F_0 \in T_{\pi_0(\Phi)}^* Q_{0,\text{free}}$ characterised in Lemma 8.14 is the *reduced Lagrangian force* associated to $\tau \oplus \mathbf{f}$. •

Note that a Lagrangian force and a reduced Lagrangian force reside in the cotangent space. This is geometrically consistent with the physical notion that a force is an object that does work on motions of a system.

We can also characterise interconnection and constraint torque-forces.

8.16 Lemma: (Lagrangian representations of interconnection and constraint forces) Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold Q . Suppose that we have velocity constraints giving rise to a smooth locally constant rank constraint distribution D . Let $\tau \oplus \mathbf{f}$ be a primary torque-force on \mathcal{B} at $\Phi \in Q$ with F and F_0 the associated Lagrangian force and reduced Lagrangian force, respectively. Then the following statements hold:

- (i) $\tau \oplus \mathbf{f}$ is an interconnection torque-force if and only if $F \in \text{ann}(T_{\Phi} Q)$;
- (ii) $\tau \oplus \mathbf{f}$ is an interconnection torque-force if and only if $F_0 \in \text{ann}(T_{\pi_0(\Phi)} Q_0)$;
- (iii) $\tau \oplus \mathbf{f}$ is a D -constraint torque-force if and only if $F \in \text{ann}(D_{\Phi})$;
- (iv) $\tau \oplus \mathbf{f}$ is a D -constraint torque-force if and only if $F_0 \in \text{ann}(D_{0,\pi_0(\Phi)})$.

Proof: (i) Let $(\mathbf{A}, \mathbf{v}) \in T_{\Phi} Q$ and let $\psi: [-r, r] \rightarrow Q$ be a differentiable motion for which $\psi'(0) = (\mathbf{A}, \mathbf{v})$. Then $\tau \oplus \mathbf{f}$ is an interconnection torque-force if and only if

$$\langle F; \psi'(0) \rangle = \sum_{a=1}^m (g(\tau_a, \omega_{\psi_a}(0)) + g(f_a, v_{\psi_a}(0))) = 0,$$

giving the result.

(ii) Let $(\mathbf{A}, \mathbf{v}) \in T_{\Phi} Q$ and let $\psi_0: [-r, r] \rightarrow Q_0$ be a differentiable reduced motion for which $\psi_0'(0) = T_{\Phi} \pi_0(\mathbf{A}, \mathbf{v})$. Then we have that $\tau \oplus \mathbf{f}$ is an interconnection torque-force if and only if

$$\langle F_0; \psi_0'(0) \rangle = \langle F; (\mathbf{A}, \mathbf{v}) \rangle = 0,$$

using the result of part (i).

The proofs of parts (iii) and (iv) proceed like those for parts (i) and (ii), asking the motions ψ and ψ_0 to satisfy $\psi'(0) \in D_{\Phi}$ and $\psi_0'(0) \in D_{0,\Phi}$, respectively. ■

Of course, the preceding constructions are immediately generalised to the case when $\boldsymbol{\tau} \oplus \mathbf{f}$ is a torque-force field, and not just a primary torque-force. Thus, if $\boldsymbol{\tau} \oplus \mathbf{f}: \mathbb{T}' \times \mathbb{T}\mathbb{Q} \rightarrow \mathbb{V}^m \oplus \mathbb{V}^m$ is a torque-force field, then we define $F: \mathbb{T}' \times \mathbb{T}\mathbb{Q} \rightarrow \mathbb{T}^*\mathbb{Q}$ and $F_0: \mathbb{T}' \times \mathbb{T}\mathbb{Q}_0 \rightarrow \mathbb{T}^*\mathbb{Q}_0$ by asking that $F(t, (\boldsymbol{\Phi}, (\mathbf{A}, \mathbf{v})))$ be the Lagrangian force and $F_0(t, (\boldsymbol{\Phi}, (\mathbf{A}, \mathbf{v})))$ be the reduced Lagrangian force associated to $\boldsymbol{\tau} \oplus \mathbf{f}(t, (\boldsymbol{\Phi}, (\mathbf{A}, \mathbf{v})))$, respectively. We call F and F_0 the *Lagrangian force field* and the *reduced Lagrangian force field*, respectively.

8.6. Equivalence of Newton–Euler and Euler–Lagrange equations. We are now in a position to state the equivalence of the Newton–Euler and Euler–Lagrange equations, completing the path from the physical world of mechanics to the mathematical world of Riemannian geometry, and so to many of the mathematical representations of mechanics that comprise “geometric mechanics.” We include variational characterisations in our list of equivalent characterisations of the solutions to the Newton–Euler equations. To this end, for a collection $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ of rigid bodies, we define

$$\begin{aligned} L_{\mathcal{B}}: \mathbb{T}\mathbb{Q}_{\text{free}} &\rightarrow \mathbb{R} \\ (\boldsymbol{\Phi}, (\mathbf{A}, \mathbf{v})) &\mapsto \frac{1}{2} \mathbf{G}_{\mathcal{B}}((\mathbf{A}, \mathbf{v}), (\mathbf{A}, \mathbf{v})), \end{aligned}$$

i.e., the kinetic energy of the free system.

With this notation, we have the following result.

8.17 Proposition: (Lagrange–d’Alembert Principle (Newton–Euler version)) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathbb{Q} . Suppose that we have velocity constraints giving rise to a smooth locally constant rank constraint distribution \mathbb{D} . Let $\boldsymbol{\tau} \oplus \mathbf{f}$ be a smooth torque-force field with F and F_0 the associated Lagrangian force and reduced Lagrangian force fields, respectively. Then, for a C^2 -arc $\phi: [t_0, t_1] \rightarrow \mathbb{Q}$, the following statements are equivalent:*

- (i) ϕ satisfies the Newton–Euler equations;
- (ii) ϕ satisfies the **Lagrange–d’Alembert principle**: for every fixed endpoint variation $\boldsymbol{\sigma}: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{Q}_{\text{free}}$ of ϕ satisfying $\delta\boldsymbol{\sigma}(t) \in \mathbb{D}_{\phi(t)}$ for every $t \in [t_0, t_1]$,

$$\delta_{\boldsymbol{\sigma}} L_{\mathcal{B}} + \sum_{a=1}^m \int_{t_0}^{t_1} (\mathbf{g}(\boldsymbol{\tau}_a(t, \boldsymbol{\phi}'(t)), \delta R_{\sigma_a}(\widetilde{t}) \circ R_{\sigma_a}^T(t)) + \mathbf{g}(f_a(t, \boldsymbol{\phi}'(t)), \delta r_{\sigma_a}(t))) dt = 0,$$

using the notation of Section 6.5, and abbreviating evaluation at $(0, t)$ with evaluation at t .

Proof: Let $\boldsymbol{\sigma}: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{Q}_{\text{free}}$ be a fixed endpoint variation of ϕ satisfying $\delta\boldsymbol{\sigma}(t) \in \mathbb{D}_{\phi(t)}$ for every $t \in [t_0, t_1]$. By the calculations of the proof of Lemma 8.12, we have

$$\begin{aligned} \frac{d}{ds} \Big|_{s=0} \int_{t_0}^{t_1} L_{\mathcal{B}}(\boldsymbol{\sigma}'(s, t)) dt &= - \sum_{a=1}^m \int_{t_0}^{t_1} (\mathbf{G}(\frac{d}{dt}(\mathbb{I}_{c,a}(\Omega_{\sigma_a}(t))) + \Omega_{\sigma_a}(t) \times (\mathbb{I}_{c,a}(\Omega_{\sigma_a}(t))), \Xi_{\sigma_a}(t)) \\ &\quad + \mathbf{G}(\frac{d}{dt}(mV_{\sigma_a}(t)) + \Omega_{\sigma_a}(t) \times V_{\sigma_a}(t), X_{\sigma_a}(t))) dt, \end{aligned}$$

where Ξ_{σ_a} and X_{σ_a} , $a \in \{1, \dots, m\}$, are defined as in the proof of Lemma 8.12:

$$\Xi_{\sigma_a}(t) = R_{\sigma_a}^T(t) \widetilde{\delta R_{\sigma_a}(t)}, \quad X_{\sigma_a}(t) = R_{\sigma_a}^T(t) (\delta r_{\sigma_a}(t)).$$

By the calculations of the proof of Theorem 7.4, namely those involved in the proof of the implication (iv) \implies (v), we have

$$\begin{aligned} \frac{d}{ds} \Big|_{s=0} \int_{t_0}^{t_1} L_{\mathcal{B}}(\boldsymbol{\sigma}'(s, t)) dt \\ = - \sum_{a=1}^m \int_{t_0}^{t_1} (\mathbb{G}(R_{\sigma_a}^T(t) (\dot{\mu}_{\sigma_a}(t)), \Xi_{\sigma_a}(t)) + \mathbb{G}(R_{\sigma_a}^T(t) (\dot{p}_{\sigma_a}(t)), X_{\sigma_a}(t))) dt. \end{aligned}$$

Now we calculate

$$\begin{aligned} \mathbb{G}(R_{\sigma_a}^T(t) (\dot{\mu}_{\sigma_a}(t)), \Xi_{\sigma_a}(t)) &= \mathbb{g}(\dot{\mu}_{\sigma_a}(t), R_{\sigma_a}(t) (R_{\sigma_a}^T(t) \widetilde{\delta R_{\sigma_a}(t)})) \\ &= \mathbb{g}(\dot{\mu}_{\sigma_a}(t), \delta R_{\sigma_a}(t) \circ R_{\sigma_a}^T(t)) \end{aligned}$$

and

$$\mathbb{G}(R_{\sigma_a}^T(t) (\dot{p}_{\sigma_a}(t)), X_{\sigma_a}(t)) = \mathbb{g}(\dot{p}_{\sigma_a}(t), \delta r_{\sigma_a}(t)),$$

where we have used Lemma 2.17 in the first calculation. Putting this all together shows that the Lagrange–d’Alembert Principle holds if and only if

$$\begin{aligned} \sum_{a=1}^m \int_{t_0}^{t_1} (\mathbb{g}(-\dot{\mu}_{\sigma_a}(t) + \tau_a(t, \boldsymbol{\phi}(t)), \delta R_{\sigma_a}(t) \circ R_{\sigma_a}^T(t)) \\ + \mathbb{g}(-\dot{p}_{\sigma_a}(t) + f_a(t, \boldsymbol{\phi}(t)), \delta r_{\sigma_a}(t))) dt = 0 \end{aligned}$$

for every fixed endpoint variation $\boldsymbol{\sigma}$ of $\boldsymbol{\phi}$. This, however, is equivalent to the instantaneous Lagrange–d’Alembert Principle of Proposition 7.6 by Lemma 8.3, and so the result follows, also by Proposition 7.6. \blacksquare

We are now in a position to state the various equivalent ways of characterising the Newton–Euler equations.

8.18 Theorem: (Equivalence of Newton–Euler and Euler–Lagrange equations)

Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathcal{Q} . Suppose that we have velocity constraints giving rise to a smooth locally constant rank constraint distribution \mathbb{D} and reduced constraint distribution \mathbb{D}_0 . Let $\boldsymbol{\tau} \oplus \mathbf{f}$ be a smooth torque-force field with F and F_0 the associated Lagrangian force and reduced Lagrangian force fields, respectively. Then, for a C^2 -arc $\boldsymbol{\phi}: [t_0, t_1] \rightarrow \mathcal{Q}$, the following statements are equivalent:

- (i) $\boldsymbol{\phi}$ satisfies the Newton–Euler equations;
- (ii) $\boldsymbol{\phi}$ satisfies the instantaneous Lagrange–d’Alembert Principle of Proposition 7.6;
- (iii) $\boldsymbol{\phi}$ satisfies the Lagrange–d’Alembert Principle of Proposition 8.17.

Now consider the associated abstract mechanical system $(\mathbb{Q}_{0,\text{free}}, \mathbb{Q}_0, \mathbb{G}_{0,\mathcal{B}}, [t_0, t_1], F_0, D_0)$. Then the preceding three statements are equivalent to the following three for the C^2 -arc $\phi_0 \triangleq \pi_0 \circ \phi$:

- (iv) ϕ_0 satisfies the forced and constrained geodesic equations;
- (v) ϕ_0 satisfies the abstract instantaneous Lagrange–d’Alembert Principle of Proposition 8.9;
- (vi) ϕ_0 satisfies the abstract Lagrange–d’Alembert Principle of Proposition 8.9.

Now consider the associated restricted abstract mechanical system $(\mathbb{Q}_0, \mathbb{G}_0, [t_0, t_1], \bar{F}_0, D_0)$, where \bar{F}_0 is the $\mathbb{G}_{0,\mathcal{B}}^{-1}$ -orthogonal projection of F_0 onto $\mathbb{T}^*\mathbb{Q}_0$. Then the preceding six statements are equivalent to the following four for the C^2 -arc ϕ_0 :

- (vii) ϕ_0 satisfies the restricted forced and constrained geodesic equations;
- (viii) ϕ_0 satisfies the restricted abstract instantaneous Lagrange–d’Alembert Principle of Proposition 8.10;
- (ix) ϕ_0 satisfies the restricted abstract Lagrange–d’Alembert Principle of Proposition 8.10;
- (x) ϕ_0 satisfies the equation

$$\overset{D}{\nabla}_{\phi'_0(t)} \phi'_0(t) = P_D \circ \mathbb{G}_0^\sharp \circ \bar{F}_0(t, \phi'_0(t)),$$

using the notation of Theorem 8.11, and $\phi'_0(t_0) \in D_{0,\phi_0(t_0)}$.

Proof: The equivalence of parts (i)–(iii) is proved in Propositions 7.6 and 8.17. The equivalence of parts (iv)–(x) is proved in Propositions 8.9 and 8.10, and Theorem 8.11. We shall prove the equivalence of parts (iii) and (vi). For brevity of notation, let us denote $\mathbb{H} = l(\mathcal{B})$.

(iii) \implies (vi) We suppose that we have a smooth connection on the bundle $\pi_0: \mathbb{Q} \rightarrow \mathbb{Q}_0$ by [Kobayashi and Nomizu 1963, Theorem II.2.1]. Suppose that $\phi: [t_0, t_1] \rightarrow \mathbb{Q}$ satisfies the Lagrange–d’Alembert Principle in the form of Proposition 8.9. Let $\sigma_0: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{Q}_0$ be a fixed endpoint variation of ϕ_0 satisfying $\delta\sigma_0(t) \in D_{\phi_0(t)}$ for every $t \in [t_0, t_1]$. Define $\bar{\sigma}: [-r, r] \times [t_0, t_1] \rightarrow \mathbb{Q}$ by asking that $t \mapsto \bar{\sigma}(s, t)$ be the horizontal lift of $t \mapsto \sigma_0(s, t)$ through $\phi(t_0)$ for each $s \in [-r, r]$. Since horizontal lift is obtained by solving differential equations—cf. the proof of [Kobayashi and Nomizu 1963, Proposition II.3.1]—and since solutions of differential equations depend regularly on initial condition, one can easily show that $\bar{\sigma}$ is of class C^2 . Let $\psi_1: [t_0, t_1] \rightarrow \mathbb{H}$ be a C^2 -curve such that $\alpha_{\mathcal{B}}(\bar{\sigma}(0, t), \psi_1(t)) = \phi(t)$ and such that $\psi_1(t_0) = \mathbf{id}$ (recalling that $\alpha_{\mathcal{B}}$ denotes the action of \mathbb{H} on \mathbb{Q}). This is possible since $\pi_0 \circ \bar{\sigma}(0, t) = \pi_0 \circ \phi(t)$ and by regularity of solutions of differential equations with respect to initial conditions. Let $\psi_2: [-r, r] \rightarrow \mathbb{H}$ be a C^2 -curve in \mathbb{H} such that $\alpha_{\mathcal{B}}(\bar{\sigma}(s, t_1), \psi_1(t_1) \circ \psi_2(s)) = \phi(t_1)$ and such that $\psi_2(0) = \mathbf{id}$. Such a curve with the prescribed regularity exists, again by regularity of solutions to differential equations with respect to initial conditions. Now define

$$\begin{aligned} \sigma: [-r, r] \times [t_0, t_1] &\rightarrow \mathbb{Q} \\ (s, t) &\mapsto \alpha_{\mathcal{B}}(\bar{\sigma}(s, t), \psi_1(t) \circ \psi_2(\frac{t-t_0}{t_1-t_0}s)). \end{aligned}$$

We can then directly verify the following properties of σ by applying the computations (8.9) and (8.10) from the proof of Theorem 8.13 to each component of \mathbb{Q} :

1. σ is of class C^2 ;

2. $\sigma(s, t_0) = \phi(t_0)$ and $\sigma(s, t_1) = \phi(t_1)$, $s \in [-r, r]$;
3. $\pi_0 \circ \sigma(s, t) = \sigma_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$;
4. $T_{\sigma(s,t)}\pi_0(\sigma'(s, t)) = \sigma'_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$;
5. $T_{\phi(t)}\pi_0(\delta\phi(t)) = \delta\sigma_0(t)$, $t \in [t_0, t_1]$.

We claim that we additionally have

6. $\delta\sigma(t) \in D_{\phi(t)}$, $t \in [t_0, t_1]$.

Indeed, by (8.10) (applied to each component of \mathbf{Q}), we have that $T_{\phi(t)}\pi_0(\delta\sigma(t)) \in D_{0,\phi(t)}$, $t \in [t_0, t_1]$. Since $\ker(T\pi_0) \subseteq D$, a moments thought then gives the desired conclusion. We then have $L_{\mathcal{B}}(\sigma'(s, t)) = L_{\mathbf{G}_0}(\sigma'_0(s, t))$ (by definition of \mathbf{G}_0) which gives

$$\left. \frac{d}{ds} \right|_{s=0} \int_{t_0}^{t_1} L_{\mathcal{B}}(\sigma'(s, t)) dt = \left. \frac{d}{ds} \right|_{s=0} \int_{t_0}^{t_1} L_{\mathbf{G}_0}(\sigma'_0(s, t)) dt. \quad (8.12)$$

Now denote

$$\delta\sigma(t) = ((\delta R_{\sigma_1}(t), \dots, \delta R_{\sigma_m}(t)), (\delta r_{\sigma_1}(t), \dots, \delta r_{\sigma_m}(t))),$$

using the notation of Section 6.5. By Lemma 8.14 we have

$$\begin{aligned} & \int_{t_0}^{t_1} \langle F_0(t, \phi(t)); \delta\sigma_0(t) \rangle dt \\ &= \sum_{a=1}^m \int_{t_0}^{t_1} (g(\tau_a(t, \phi'(t)), \delta R_{\sigma_a}(t) \widetilde{\circ} R_{\sigma_a}^T(t)) + g(f_a(t, \phi'(t)), \delta r_{\sigma_a}(t))) dt. \end{aligned} \quad (8.13)$$

The result follows immediately by combining (8.12) and (8.13).

(vi) \implies (iii) Suppose that ϕ_0 satisfies the abstract Lagrange–d'Alembert Principle and let $\sigma: [-r, r] \times [t_0, t_1] \rightarrow \mathbf{Q}$ be a fixed endpoint variation of ϕ satisfying $\delta\sigma(t) \in D_{\phi(t)}$ for each $t \in [t_0, t_1]$. Define $\sigma_0 = \pi_0 \circ \sigma$. Then one readily verifies that:

1. σ_0 is of class C^2 ;
2. $\sigma_0(s, t_0) = \phi_0(t_0)$ and $\sigma_0(s, t_1) = \phi_0(t_1)$, $s \in [-r, r]$;
3. $\pi_0 \circ \sigma(s, t) = \sigma_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$;
4. $T_{\sigma(s,t)}\pi_0(\sigma'(s, t)) = \sigma'_0(s, t)$, $(s, t) \in [-r, r] \times [t_0, t_1]$;
5. $\delta\sigma_0(t) \in D_{0,\phi_0(t)}$, $t \in [t_0, t_1]$.

Then, just as in the first part of the proof, the equations (8.12) and (8.13) obtain, and this part of the result follows. \blacksquare

Of course, we have the following immediate corollary which makes the usual connection between the Newton–Euler and the Euler–Lagrange equations in the absence of forces and constraints.

8.19 Corollary: (Solutions of the unforced unconstrained Newton–Euler equations as geodesics (multibody version)) *Consider an interconnected rigid body system with \mathcal{S} a Newtonian space model, $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_m)$ rigid bodies in body reference spaces $\mathcal{B}_1, \dots, \mathcal{B}_m$, and configuration manifold \mathbf{Q} . Let \mathbb{T} be a time axis with $\mathbb{T}' \subseteq \mathbb{T}$ a time interval. Then, for a C^2 -motion $\phi: \mathbb{T}' \rightarrow \mathbf{Q}$ with reduced motion $\phi_0 = \pi_0 \circ \phi$, the following statements are equivalent:*

- (i) ϕ satisfies the unforced Newton–Euler equations;
- (ii) for every compact $\mathbb{T}'' \subseteq \mathbb{T}'$, $\phi_0|_{\mathbb{T}''}$ satisfies the Euler–Lagrange equations for $L_{\mathbb{G}_0}$, i.e., $\phi_0|_{\mathbb{T}''}$ is an $L_{\mathbb{G}_0}$ -extremal;
- (iii) ϕ_0 is a geodesic for the Riemannian metric \mathbb{G}_0 .

References

- Abraham, R. and Marsden, J. E. [1978] *Foundations of Mechanics*, 2nd edition, Addison Wesley: Reading, MA, ISBN: 978-0-8218-4438-0.
- Abraham, R., Marsden, J. E., and Ratiu, T. S. [1988] *Manifolds, Tensor Analysis, and Applications*, number 75 in Applied Mathematical Sciences, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-96790-5.
- Arnol'd, V. I. [1978] *Mathematical Methods of Classical Mechanics*, translated by K. Vogtmann and A. Weinstein, number 60 in Graduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 0-387-96890-3, New edition: [Arnol'd 1989].
- [1989] *Mathematical Methods of Classical Mechanics*, translated by K. Vogtmann and A. Weinstein, 2nd edition, number 60 in Graduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-96890-2, First edition: [Arnol'd 1978].
- Artz, R. E. [1981] *Classical mechanics in Galilean space-time*, Foundations of Physics, An International Journal Devoted to the Conceptual Bases and Fundamental Theories of Modern Physics, **11**(9-10), pages 679–697, ISSN: 0015-9018, DOI: [10.1007/BF00726944](https://doi.org/10.1007/BF00726944).
- Berger, M. [1987] *Geometry I*, Universitext, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-540-11658-5.
- Bhand, A. and Lewis, A. D. [2005] *Rigid body mechanics in Galilean spacetimes*, Journal of Mathematical Physics, **46**(10), 29 pages, page 102902, ISSN: 0022-2488, DOI: [10.1063/1.2060547](https://doi.org/10.1063/1.2060547).
- Bloch, A. M. [2003] *Nonholonomic Mechanics and Control*, number 24 in Interdisciplinary Applied Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-95535-3.
- Brogliato, B. [1999] *Nonsmooth Mechanics, Models, Dynamics and Control*, Communications and Control Engineering Series, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-1-85233-143-6.
- Bullo, F. and Lewis, A. D. [2004] *Geometric Control of Mechanical Systems, Modeling, Analysis, and Design for Simple Mechanical Systems*, number 49 in Texts in Applied Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-22195-3.
- Cendra, H., Holm, D. D., Marsden, J. E., and Ratiu, T. S. [1998] *Lagrangian reduction, the Euler–Poincaré equations, and semidirect products*, American Mathematical Society Translations, Series 2, **186**, pages 1–25, ISSN: 0065-9290.
- Cohn, D. L. [2013] *Measure Theory*, 2nd edition, Birkhäuser Advanced Texts, Birkhäuser: Boston/Basel/Stuttgart, ISBN: 978-1-4614-6955-1.
- Cortés, J., de León, M., Martín de Diego, D., and Martínez, S. [2001] *Mechanical systems subjected to generalized non-holonomic constraints*, Proceedings of the Royal Society. London. Series A. Mathematical and Physical Sciences, **457**(2007), pages 651–670, ISSN: 1364-5021, DOI: [10.1098/rspa.2000.0686](https://doi.org/10.1098/rspa.2000.0686).

- Crampin, M. [1986] *On the concept of angular velocity*, European Journal of Physics, **7**(4), pages 287–293, ISSN: 0143-0807, DOI: [10.1088/0143-0807/7/4/014](https://doi.org/10.1088/0143-0807/7/4/014).
- Flannery, M. R. [2005] *The enigma of nonholonomic constraints*, American Journal of Physics, **73**(3), pages 265–272, ISSN: 0002-9505, DOI: [10.1119/1.1830501](https://doi.org/10.1119/1.1830501).
- Glocker, C. [2001] *Set Valued Force Laws, Dynamics of Non-smooth Systems*, number 1 in Lecture Notes in Applied Mechanics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-642-53595-6.
- Goldstein, H. [1951] *Classical Mechanics*, Addison Wesley: Reading, MA, New edition: [Goldstein, Poole, Jr, and Safko 2001].
- Goldstein, H., Poole, Jr, C. P., and Safko, J. L. [2001] *Classical Mechanics*, 3rd edition, Addison Wesley: Reading, MA, ISBN: 978-0-201-65702-9, Original edition: [Goldstein 1951].
- Gràcia, X., Marin-Solano, J., and Muñoz-Lecanda, M. C. [2003] *Some geometric aspects of variational calculus in constrained systems*, Reports on Mathematical Physics, **51**(1), pages 127–148, ISSN: 0034-4877, DOI: [10.1016/S0034-4877\(03\)80006-X](https://doi.org/10.1016/S0034-4877(03)80006-X).
- Hatcher, A. [2002] *Algebraic Topology*, Cambridge University Press: New York/Port Chester/Melbourne/Sydney, ISBN: 978-0-521-79540-1.
- Husemoller, D. [1994] *Fibre Bundles*, 3rd edition, number 20 in Graduate Texts in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-94087-8.
- Jafarpour, S. and Lewis, A. D. [2014] *Time-Varying Vector Fields and Their Flows*, Springer Briefs in Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-319-10138-5.
- Kanno, Y. [2011] *Nonsmooth Mechanics and Convex Optimization*, CRC Press: Boca Raton, FL, ISBN: 978-1-4200-9423-7.
- Kharlomov, P. V. [1992] *A critique of some mathematical models of mechanical systems with differential constraints*, Journal of Applied Mathematics and Mechanics, Translation of *Rossiiskaya Akademiya Nauk. Prikladnaya Matematika i Mekhanika*, **56**(4), pages 584–594, ISSN: 0021-8928, DOI: [10.1016/0021-8928\(92\)90016-2](https://doi.org/10.1016/0021-8928(92)90016-2).
- Kobayashi, S. and Nomizu, K. [1963] *Foundations of Differential Geometry*, volume 1, number 15 in Interscience Tracts in Pure and Applied Mathematics, Interscience Publishers: New York, NY, Reprint: [Kobayashi and Nomizu 1996].
- [1996] *Foundations of Differential Geometry*, volume 1, Wiley Classics Library, John Wiley and Sons: New York, NY, ISBN: 978-0-471-15733-5, Original: [Kobayashi and Nomizu 1963].
- Kozlov, V. V. [1992] *The problem of realizing constraints in dynamics*, Journal of Applied Mathematics and Mechanics, Translation of *Rossiiskaya Akademiya Nauk. Prikladnaya Matematika i Mekhanika*, **56**(4), pages 594–600, ISSN: 0021-8928, DOI: [10.1016/0021-8928\(92\)90017-3](https://doi.org/10.1016/0021-8928(92)90017-3).
- Lagrange, J. L. [1788] *Mécanique analytique*, Chez la Veuve Desaint: Paris, Translation: [Lagrange 1997].
- [1997] *Analytical Mechanics*, translated by A. Boissonnade and V. N. Vagliente, number 191 in Boston Studies in the Philosophy of Science, Kluwer Academic Publishers: Dordrecht, ISBN: 978-0-7923-4349-3, Original: [Lagrange 1788].
- Lewis, A. D. [1996] *The geometry of the Gibbs–Appell equations and Gauss’s Principle of Least Constraint*, Reports on Mathematical Physics, **38**(1), pages 11–28, ISSN: 0034-4877, DOI: [10.1016/0034-4877\(96\)87675-0](https://doi.org/10.1016/0034-4877(96)87675-0).

- Lewis, A. D. [1998] *Affine connections and distributions with applications to nonholonomic mechanics*, Reports on Mathematical Physics, **42**(1/2), pages 135–164, ISSN: 0034-4877, DOI: [10.1016/S0034-4877\(98\)80008-6](https://doi.org/10.1016/S0034-4877(98)80008-6).
- [2007] *Is it worth learning differential geometric methods for modelling and control of mechanical systems?*, Robotica, **26**(6), pages 765–777, ISSN: 0263-5747, DOI: [10.1017/S0263574707003815](https://doi.org/10.1017/S0263574707003815).
- Lewis, A. D. and Murray, R. M. [1995] *Variational principles for constrained systems: Theory and experiment*, International Journal of Non-Linear Mechanics, **30**(6), pages 793–815, ISSN: 0020-7462, DOI: [10.1016/0020-7462\(95\)00024-0](https://doi.org/10.1016/0020-7462(95)00024-0).
- Libermann, P. and Marle, C.-M. [1987] *Symplectic Geometry and Analytical Mechanics*, number 35 in Mathematics and its Applications, D. Reidel Publishing Company: Dordrecht/Boston/Lancaster/Tokyo, ISBN: 978-90-2772-438-0.
- Marsden, J. E. and Hughes, T. R. [1983] *Mathematical Foundations of Elasticity*, Prentice-Hall: Englewood Cliffs, NJ, ISBN: 0-486-67865-2, Reprint: [Marsden and Hughes 1994].
- [1994] *Mathematical Foundations of Elasticity*, Dover Publications, Inc.: New York, NY, ISBN: 978-0-486-67865-8, Original: [Marsden and Hughes 1983].
- Marsden, J. E. and Scheurle, J. [1993] *The reduced Euler–Lagrange equations*, in *Dynamics and Control of Mechanical Systems, The Falling Cat and Related Problems*, Fields Institute Workshop, (Waterloo, Canada, Mar. 1992), edited by M. J. Enos, 1 Fields Institute Communications, pages 139–164, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-8218-9200-8.
- Monteiro Marques, M. D. P. [1993] *Differential Inclusions in Nonsmooth Mechanical Problems, Shocks and Dry Friction*, number 9 in Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser: Boston/Basel/Stuttgart, ISBN: 978-3-7643-2900-6.
- Murray, R. M., Li, Z. X., and Sastry, S. S. [1994] *A Mathematical Introduction to Robotic Manipulation*, CRC Press: Boca Raton, FL, ISBN: 978-0-8493-7981-9.
- O’Reilly, O. M. [2008] *Intermediate Dynamics for Engineers, A Unified Treatment of Newton–Euler and Lagrangian Mechanics*, Cambridge University Press: New York/Port Chester/Melbourne/Sydney, ISBN: 978-0-521-87483-0.
- Papastavridis, J. G. [1999] *Tensor Calculus and Analytical Dynamics*, Library of Engineering Mathematics, CRC Press: Boca Raton, FL, ISBN: 978-0-8493-8514-8.
- Pars, L. A. [1965] *A Treatise on Analytical Dynamics*, John Wiley and Sons: New York, NY, ISBN: 978-0-918024-07-7.
- Spivak, M. [2010] *Physics for Mathematicians, Mechanics I*, Publish or Perish, Inc.: Houston, TX, ISBN: 978-0-914098-32-4.
- Truesdell, C. and Noll, W. [1965] *The Non-Linear Field Theories of Mechanics*, number 3 in Handbuch der Physik, Springer-Verlag: New York/Heidelberg/Berlin, New edition: [Truesdell and Noll 2004].
- [2004] *The Non-Linear Field Theories of Mechanics*, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-3-540-02779-9, Original: [Truesdell and Noll 1965].
- Whittaker, E. T. [1937] *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge University Press: New York/Port Chester/Melbourne/Sydney, Reprint: [Whittaker 1988].
- [1988] *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, with a foreword by W. McCrea, Cambridge Mathematical Library, Cambridge University Press:

New York/Port Chester/Melbourne/Sydney, ISBN: 978-0-521-35883-5, Original: [[Whitaker 1937](#)].