Slides for MATH/MTHE 335, Winter 2020

Andrew D. Lewis

Department of Mathematics and Statistics, Queen's University

Queen's UNIVERSITY

March 7, 2022

What follows is some sort of transcription of what I put on the board every day during the winter term of 2020. These notes are based on LaTeX lecture notes prepared by Joe Grosso in real time during lectures. Beware of typos.

Andrew Lewis

Lecture 1
General System Theory

Definition

A ***general input-output system*** is a triple $(\mathscr{U}, \mathscr{Y}, \mathscr{B})$, where

(i) $\mathscr{U}$ is a set (the set of ***inputs***),

(ii) $\mathscr{Y}$ is a set (the set of ***outputs***), and

(iii) $\mathscr{B} \subseteq \mathscr{U} \times \mathscr{Y}$ (the ***behaviours***).

- An element of $\mathscr{B}$ looks like (μ, η) with μ an input and η an output. If $\mu \in \mathscr{U}$, we denote

$$\mathscr{B}(\mu) = \{\eta \in \mathscr{Y} \mid (\mu, \eta) \in \mathscr{B}\}.$$

This is the set of all outputs corresponding to the input $\mu \in \mathscr{U}$.

Definition

A ***functional input/output system*** is a general input/output system (U(,)$\mathscr{Y}, \mathscr{B}$) such that $\mathscr{B}$ is the graph of some mapping $\Phi\colon \mathscr{U} \to \mathscr{Y}$:

$$\mathscr{B} = \mathrm{graph}(\Phi) = \{(\mu, \Phi(\mu)) \mid \mu \in \mathscr{U}\}.$$

States

Definition

A ***response function*** for a general input/output system $(\mathscr{U}, \mathscr{Y}, \mathscr{B})$ is a mapping $\rho\colon \mathscr{U} \times X \to \mathscr{Y}$ such that

$$\mathscr{B} = \{(\mu, \rho(\mu, x)) \mid \mu \in \mathscr{U},\ x \in X\}.$$

The set X is called a ***state object***.

- One should think of X as parameterising the outputs for a fixed input.

Definition

A ***linear general input/output system*** is a general input/output system $(\mathscr{U}, \mathscr{Y}, \mathscr{B})$ such that

(i) $\mathscr{U}$ and $\mathscr{Y}$ are vector spaces and

(ii) $\mathscr{B} \subseteq \mathscr{U} \oplus \mathscr{Y}$ is a subspace.

States

- One also has linear response functions, meaning that $\mathscr{U}$, $\mathscr{Y}$, and X are vector spaces, and $\rho\colon \mathscr{U} \oplus \mathsf{X} \to \mathscr{Y}$ is linear.
- *Nontrivial fact:* If $(\mathscr{U}, \mathscr{Y}, \mathscr{B})$ is a linear input/output system it has a linear response function.

Reading for Lecture 1

Material related to this lecture can be found in the following sections of the course notes:

1. Sections V-2.1.2, V-2.1.3, and V-2.1.5.

Lecture 2
Time and set-valued functions of time

- For a (discrete or continuous time-domain) $\mathbb{T}$, a ***sub-time-domain*** is a subset $\mathbb{T}' = \mathbb{T} \cap I$ where I is an interval.
- For a time-domain $\mathbb{T}$ and a set X, recall that

$$X^{\mathbb{T}} = \{f \colon \mathbb{T} \to X\}$$

  denotes the set of all X-valued functions on $\mathbb{T}$.
- We shall also require "partial functions," defined by

$$X^{(\mathbb{T})} = \{f \colon \mathbb{T}' \to X \mid \ \mathbb{T}' \subseteq \mathbb{T} \text{ is a sub-time-domain}\}.$$

- Call $\mathbb{T}'$ the ***domain*** of f, denoted $\mathrm{dom}(f)$.
- Why do we need partial functions?

Time and set-valued functions of time

Example

Consider a general input/output system with

$$\mathscr{U} = \mathsf{L}^1(\mathbb{R}; \mathbb{R}), \quad \mathscr{Y} = \mathsf{C}^0(\mathbb{R}, \mathbb{R}).$$

Given $\mu \in \mathscr{U}$, the outputs corresponding to this input are given by solutions to the following initial value problem:

$$\dot{\eta}(t) = \mu(t)\eta^2(t), \quad \eta(0) = y_0.$$

This is a differential equation, which we can solve by the method of separation to give

$$\eta(t) = \frac{y_0}{1 - y_0 \int_{t_0}^{t} \mu(\tau)\,\mathrm{d}\tau}.$$

Time and set-valued functions of time

Example (cont'd)

For example, when $\mu(t) = u_0$, i.e., μ is constant (leave aside for the moment that such inputs are not in L^1; they are in L^1_{loc} which we have yet to define), then

$$\eta(t) = \frac{y_0}{1 - y_0\mu_0(t - t_0)}.$$

Punchline: Although the differential equation defining the outputs seems nice enough, and although the inputs are nice, the outputs can—and often do—blow up in finite time.

General time systems

Definition

A ***general time-system*** is $(U, Y, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathscr{B})$ such that

(i) U is a set (the set of input values),

(ii) Y is a set (the set of output values),

(iii) $\mathbb{T}$ is a time domain,

(iv) $\mathscr{U} \subseteq U^{(\mathbb{T})}$ (the inputs),

(v) $\mathscr{Y} \subseteq Y^{(\mathbb{T})}$ (the outputs),

(vi) $\mathscr{B} \subseteq \mathscr{U} \times \mathscr{Y}$ is such that, if $(\mu, \eta) \in B$, then μ and η have the same domain.

- Thus inputs are things like $\mu\colon \mathbb{T}' \to U$, where $\mathbb{T}'$ is a sub-time-domain of $\mathbb{T}$.
- Similarly, outputs are things like $\eta\colon \mathbb{T}' \to Y$, where $\mathbb{T}'$ is a sub-time-domain of $\mathbb{T}$.

General time systems

- There is a notational convention for general time systems emerging that will be repeated throughout the course:
  1. Sets are denote by uppercase roman letters like U, Y, and X.
  2. Sets of functions of time with values in these sets are denoted by uppercase script letters like $\mathscr{U}$, $\mathscr{Y}$, and $\mathscr{X}$.
  3. Points in these sets are denoted by lowercase roman letters like u, y, and x.
  4. Specific functions of time with values in these sets are denoted by lowercase greek letters that are meant to be brothers of their roman brothers, e.g., μ, η, and ξ.

Attributes of general time systems

- We will examine, in a general context, three attributes of general time systems. We shall come to a better understanding of these when we talk about concrete classes of systems later on in the course. The point here is that these can be discussed in generality.
- In doing this, we suppose we are given a distinguished "starting time" $t_0 \in \mathbb{T}$.
- For $\xi \in X^{\mathbb{T}}$ and $t \geq t_0$ we have

$$\xi_{[t_0,t)} = \xi|[t_0,t) \quad \text{and} \quad \xi_{[t_0,t]} = \xi|[t_0,t].$$

- If $(U, Y, \mathbb{T}, \mathscr{U}, \mathscr{Y})$ is a general time system, we denote

$$\mathscr{B}_{[t_0,t)} = \{(\mu_{[t_0,t)}, \eta_{[t_0,t)}) \mid (\mu,\eta) \in \mathscr{B}\}.$$

  Similarly we have $\mathscr{B}_{[t_0,t]}$.
- If $\mu \in \mathscr{U}$ is a fixed input, then denote

$$\mathscr{B}(\mu)_{[t_0,t)} = \{\eta_{[t_0,t)} \mid (\mu,\eta) \in \mathscr{B}\}.$$

Attributes of general time systems

Definition

Let $(U, Y, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathscr{B})$ be a general time system. It is:

(i) ***causal*** from $t_0 \in \mathbb{T}$ if

$$\begin{aligned} &(\mu_1)_{[t_0,t]} = (\mu_2)_{[t_0,t]} \\ \implies \; &\mathscr{B}(\mu_1)_{[t_0,t]} = \mathscr{B}(\mu_2)_{[t_0,t]} \;\forall\, t \geq t_0; \end{aligned}$$

(ii) ***strongly causal*** from $t_0 \in \mathbb{T}$ if

$$\begin{aligned} &(\mu_1)_{[t_0,t)} = (\mu_2)_{[t_0,t)} \\ \implies \; &\mathscr{B}(\mu_1)_{[t_0,t]} = \mathscr{B}(\mu_2)_{[t_0,t]} \;\forall\, t \geq t_0. \end{aligned}$$

- *Causal:* The behaviours at time t are determined by inputs for times not beyond t.
- *Strongly causal:* The behaviours at time t are determined by inputs for times strictly less than t.

Attributes of general time systems

Definition

A general time system $(U, Y, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathscr{B})$ is ***finitely observable*** from $\tau \geq t_0$ if, for every input μ and for outputs $\eta_1, \eta_2 \in \mathscr{B}(\mu)$, we have

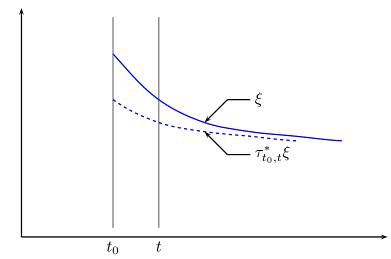
$$\begin{aligned} &(\eta_1)_{[t_0,\tau]} = (\eta_2)_{[t_0,\tau]} \\ \implies \; &(\eta_1)_{\geq t_0} = (\eta_2)_{\geq t_0}. \end{aligned}$$

- The idea is that the system gets all the information it needs to determine outputs by time τ.
- This seems like a peculiar property, but there are systems that are not finitely observable from all $\tau \geq t_0$.

Attributes of general time systems

- For $t \geq t_0$, we can shift a signal to the left by $t - t_0$:

$$\tau^*_{t_0,t}\xi(s) = \xi(s - (t - t_0)), \qquad s \geq t_0.$$



- We denote

$$\tau^*_{t_0,t}\mathscr{B} = \{(\tau^*_{t_0,t}\mu, \tau^*_{t_0,t}\eta) \mid (\mu, \eta) \in \mathscr{B}\}.$$

Attributes of general time systems

Definition

A general time system $(U, Y, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathscr{B})$ is
(i) ***stationary*** if $\tau^*_{t_0,t}\mathscr{B} \subseteq \mathscr{B}$ and is
(ii) ***strongly stationary*** if $\tau^*_{t_0,t}\mathscr{B} = \mathscr{B}$.

- *Stationary:* behaviours are "shift-invariant."
- *Strongly stationary:* behaviours are "shift-invariant" and no behaviours are lost by shifting.

Reading for Lecture 2

Material related to this lecture can be found in the following sections of the course notes:

1. Sections V-2.2.6, V-2.2.7, and V-2.2.8.

Lecture 3
More classes of signals

- Let $\mathbb{T}$ be a continuous time-domain and let $p \in [1, \infty]$. Denote by

$$\mathsf{L}^p_{\text{loc}} = \{f \colon \mathbb{T} \to \mathbb{F} \mid f|\mathbb{K} \in \mathsf{L}^p(\mathbb{K};\mathbb{F}) \text{ for every compact interval } \mathbb{K} \subseteq \mathbb{T}\}$$

Examples

1. Take $\mathbb{T} = (0, 1]$, $f(t) = t^{-1}$. Then $f \notin \mathsf{L}^p((0, 1], \mathbb{R})$ for *all* p.
   However, $f \in \mathsf{L}^p_{\text{loc}}((0, 1], \mathbb{R})$ for every $p \in [1, \infty]$. Indeed, let $\mathbb{K} \subseteq (0, 1]$ be a compact interval so that $\mathbb{K} = [a, b]$ for $0 < a < b \leq 1$. Thus, since f is continuous and $\mathbb{K}$ is compact:

$$f|\mathbb{K} \in \mathsf{L}^p(\mathbb{K};\mathbb{R}).$$

2. Let $\mathbb{T} = \mathbb{R}$ and take

$$f(t) = \begin{cases} t^{-1}, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

   Then $f \in \mathsf{L}^p_{\text{loc}}(\mathbb{R};\mathbb{R})$ for *no* $p \in [1, \infty]$. Indeed, if $\mathbb{K} \subseteq \mathbb{R}$ is compact with $0 \in \text{int}(\mathbb{K})$, then $f|\mathbb{K} \notin \mathsf{L}^p(\mathbb{K};\mathbb{R})$.

More classes of signals

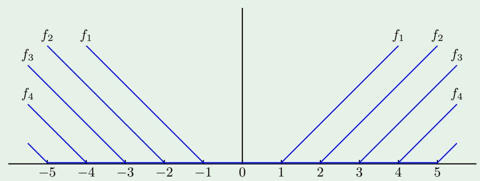
- In the spaces $\mathsf{L}^p_{\text{loc}}(\mathbb{T};\mathbb{F})$, we have a notion of convergence.

Definition

Let $\mathbb{T}$ be a continuous time-domain and let $(f_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence in $\mathsf{L}^p_{\text{loc}}(\mathbb{T};\mathbb{F})$. The sequence ***converges*** to $f \in \mathsf{L}^p_{\text{loc}}(\mathbb{T};\mathbb{F})$ if, for every compact subinterval $\mathbb{K} \subseteq \mathbb{T}$, the sequence $(f_j|\mathbb{K})_{j\in\mathbb{Z}_{>0}}$ converges to $f|\mathbb{K}$.

Example

Consider the sequence in $\mathsf{L}^p_{\text{loc}}(\mathbb{R};\mathbb{R})$ depicted here:



More classes of signals

Example (cont'd)

We claim that $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathsf{L}^p_{\text{loc}}(\mathbb{R};\mathbb{R})$ for every $p \in [1,\infty]$. Let $\mathbb{K} \subseteq \mathbb{R}$ be a compact subinterval. Choose $N \in \mathbb{Z}_{>0}$ be sufficiently large that $\mathbb{K} \subseteq [-N,N]$. Then, by definition of f_j, if $j \geq N$, $f_j|[-N,N] = 0$ which means that $f_j|\mathbb{K} = 0$. Thus $(f_j|\mathbb{K})_{j\in\mathbb{Z}_{>0}}$ clearly converges to zero in $\mathsf{L}^p(\mathbb{K};\mathbb{R})$. This, by definition, means that $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathsf{L}^p_{\text{loc}}(\mathbb{R};\mathbb{R})$.

- Similarly, we can consider $\mathsf{C}^0(\mathbb{T};\mathbb{F})$ and talk about convergence in this space.

Definition

A sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{C}^0(\mathbb{T};\mathbb{F})$ converges to $f \in \mathsf{C}^0(\mathbb{T};\mathbb{F})$ if, for every compact interval $\mathbb{K} \subseteq \mathbb{T}$, $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to $f|\mathbb{K}$ in $\mathsf{C}^0(\mathbb{K};\mathbb{F})$ using the ∞-norm.

Reading for Lecture 3

Material related to this lecture can be found in the following sections of the course notes:

1. Sections III-6.5.2 and III-6.5.4, Proposition III-6.2.11, and Section IV-1.3.5.

Lecture 4
More classes of signals (cont'd)

- There are discrete-time analogues to the classes $\mathsf{L}^p_{\mathsf{loc}}(\mathbb{T};\mathbb{F})$ and $\mathsf{C}^0(\mathbb{T};\mathbb{F})$ of continuous-time signals that we discussed last time. But they are easier.
- Let $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ be a discrete time-domain. Denote

$$\ell_{\mathsf{loc}}(\mathbb{T};\mathbb{F}) = \mathbb{F}^{\mathbb{T}}.$$

- Because, for finite discrete time-domains $\mathbb{T}$, there are no differences between the spaces $\ell^p(\mathbb{T};\mathbb{F})$, $p \in [1,\infty]$, there is no discrimination between various flavours of discrete-time signals spaces in this setting as there is with continuous-time signal spaces.

More classes of signals

- As concerns convergence in $\ell_{\mathrm{loc}}(\mathbb{T};\mathbb{F})$, the rôle of compact subintervals is played by finite subsets.

Definition

A sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\ell_{\mathrm{loc}}(\mathbb{T};\mathbb{F})$ converges to f in $\ell_{\mathrm{loc}}(\mathbb{T};\mathbb{F})$ if, for every finite $\mathbb{K}\subseteq\mathbb{T}$, the sequence $(f_j|\mathbb{K})_{j\in\mathbb{Z}_{>0}}$ converges to $f|\mathbb{K}$.

- Convergence in $\ell_{\mathrm{loc}}(\mathbb{T};\mathbb{F})$ is really just pointwise convergence. Understand this!
- In all of the above cases, convergence is not, generally, norm convergence. It is norm convergence when and only when $\mathbb{T}$ is compact (in the continuous-time case) or finite (in the discrete-time case). The convergence is with respect to seminorms. The seminorms are:
  1. $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{T};\mathbb{F})$: $\|f\|_{\mathbb{K},p}=\|f|\mathbb{K}\|$, $\mathbb{K}\subseteq\mathbb{T}$ compact;
  2. $\mathsf{C}^0(\mathbb{T};\mathbb{F})$: $\|f\|_{\mathbb{K},\infty}=\|f|\mathbb{K}\|_\infty$, $\mathbb{K}\subseteq\mathbb{T}$ compact;
  3. $\ell_{\mathrm{loc}}(\mathbb{T};\mathbb{F})$: $\|f\|_{\mathbb{K},p}=\|f|\mathbb{K}\|_p$, $\mathbb{K}\subseteq\mathbb{T}$ finite (convergence is independent of p).

More classes of signals

- Convergence in these spaces in equivalent to convergence in each of the seminorms.
- One can also talk about continuity of functions to or from these spaces using sequences, i.e., f is continuous if and only if $\lim_{j\to\infty}f(x_j)=f(x)$ for any sequence $(x_j)_{j\in\mathbb{Z}_{>0}}$ converging to x.
- *Punchline:* We can fairly easily enlarge our classes of signal spaces, using the existing classes of signal spaces we learned about in MATH/MTHE 334.

Ordinary differential equations

- In your past life, integrals were things you computed and differential equations were things you solved.
- In MATH/MTHE 334 we saw how thinking about the definition of an integral is important, not just the computation of integrals.
- Here we take a similarly elevated view of differential equations. We will distinguish between a differential equation and a solution to a differential equation.

Definition

Let $X \subseteq \mathbb{R}^n$ be open and let $\mathbb{T} \subseteq \mathbb{R}$ be a continuous time-domain. An ***ordinary differential equation*** with state space X and time-domain $\mathbb{T}$ is a mapping $\widehat{\boldsymbol{F}}\colon \mathbb{T} \times X \to \mathbb{R}^n$.

Ordinary differential equations

- Solutions of differential equations are required to have a certain property.

Definition

Let $\mathbb{T}$ be a continuous time-domain. A mapping $f\colon \mathbb{T} \to \mathbb{R}$ is ***locally absolutely continuous*** if there exists $t \in \mathbb{T}$ and $g \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T};\mathbb{R})$ such that

$$f(t) = f(t_0) + \int_{t_0}^{t} g(\tau)\mathrm{d}\tau$$

for some $t_0 \in \mathbb{T}$ (the definition in independent of this choice).

- Properties of locally absolutely continuous function f as in the definition:
  1. f is continuous;
  2. f is differentiable almost everywhere, and $f'(t) = g(t)$ for almost every $t \in \mathbb{T}$.

Ordinary differential equations

Definition

Let $\widehat{\boldsymbol{F}}: \mathbb{T} \times X \to \mathbb{R}^n$ be an ordinary differential equation. A ***solution*** to $\widehat{\boldsymbol{F}}$ is a locally absolutely continuous $\boldsymbol{\xi}: \mathbb{T}' \to X$ where $\mathbb{T}' \subseteq \mathbb{T}$ is a subinterval and where

$$\dot{\boldsymbol{\xi}}(t) = \widehat{\boldsymbol{F}}(t, \boldsymbol{\xi}(t))$$

for almost every $t \in \mathbb{T}'$.

- We will give conditions for $\widehat{\boldsymbol{F}}$ that ensure the existence of solutions and some sort of uniqueness of solutions.
- The conditions are a tiny bit complicated. But they are useful to think about carefully:
  1. because they are an important part of the theory of ordinary differential equations;
  2. because, when we talk about continuity of *systems* subsequently, we will make use of conditions like those we give here.

Ordinary differential equations

Definition

Let $X \subseteq \mathbb{R}^n$ be open and let $\boldsymbol{f}: X \to \mathbb{R}^m$. Say that

(i) Say that f is ***Lipschitz*** if there exists $L \in \mathbb{R}_{>0}$ (called a ***Lipshitz constant***) such that

$$\|\boldsymbol{f}(\boldsymbol{x}_1) - \boldsymbol{f}(\boldsymbol{x}_2)\| \leq L\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|, \qquad \boldsymbol{x}_1, \boldsymbol{x}_2 \in X;$$

(ii) $\boldsymbol{f}$ is ***locally Lipschitz*** if, for every compact $K \subseteq X$, $f|K$ is Lipschitz.

- Properties:
  1. If f is locally Lipshitz, it is continuous.
  2. If f is continuously differentiable, it is locally Lipschitz.

Examples

1. The function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \sqrt{|x|}$ is continuous but not locally Lipschitz.
2. The function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$ is locally Lipschitz (in fact, it is Lipschitz with Lipschitz constant 1) but not continuously differentiable.

Reading for Lecture 4

Material related to this lecture can be found in the following sections of the course notes:

1. Sections III-6.5.1 and IV-1.2.5.
2. Sections V-3.1.3.1 and V-3.2.1.
3. Section II-1.10.8.

Lecture 5
Ordinary differential equations (cont'd)

Theorem

Let $\mathbb{T} \subseteq \mathbb{R}$ be a continuous time-domain, let $X \subseteq \mathbb{R}^n$ be open, and let $\widehat{\boldsymbol{F}} \colon \mathbb{T} \times X \to \mathbb{R}^n$ be an ordinary differential equation. Let $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$. Suppose that:

(i) for every $t \in \mathbb{T}$, the function $\boldsymbol{x} \mapsto \widehat{\boldsymbol{F}}(t, \boldsymbol{x})$ is locally Lipschitz;

(ii) for every $\boldsymbol{x} \in X$, the mapping $t \mapsto \widehat{\boldsymbol{F}}(t, \boldsymbol{x})$ is locally integrable;

(iii) for every $(t, \boldsymbol{x}) \in \mathbb{T} \times X$, there exists $r, \rho \in \mathbb{R}_{>0}$ and $g_0, g_1 \in \mathsf{L}^1([t-\rho, t+\rho]; \mathbb{R}_{\geq 0})$ such that

(a) $\|\widehat{\boldsymbol{F}}(t', \boldsymbol{x}')\| \leq g_0(t')$, $(t', x') \in [t-\rho, t+\rho] \times \mathsf{B}(r, \boldsymbol{x})$;

(b) $\|\widehat{\boldsymbol{F}}(t', \boldsymbol{x}_1) - \widehat{\boldsymbol{F}}(t', \boldsymbol{x}_2)\| \leq g_1(t')\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$, $t' \in [t-\rho, t+\rho]$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathsf{B}(r, \boldsymbol{x})$.

Then there exists a solution $\boldsymbol{\xi} \colon \mathbb{T}' \to X$ for $\widehat{\boldsymbol{F}}$ with $t_0 \in \mathbb{T}'$ and $\boldsymbol{\xi}(t_0) = \boldsymbol{x}_0$. Moreover, if $\widetilde{\boldsymbol{\xi}} \colon \mathbb{T}'' \to X$ is another solution satisfying $\widetilde{\boldsymbol{\xi}}(t_0) = \boldsymbol{x}_0$, then

$$\boldsymbol{\xi}(t) = \widetilde{\boldsymbol{\xi}}(t), \quad t \in \mathbb{T}' \cap \mathbb{T}''.$$

Ordinary differential equations

- Generally, one cannot "solve" differential equations in any meaningful way. But one still wants to be able to talk about solutions in an organised way.
- For $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$, let $I_{\widehat{\boldsymbol{F}}}(t_0, \boldsymbol{x}_0) \subseteq \mathbb{T}$ be the largest interval on which a solution exists satisfying $\boldsymbol{\xi}(t_0) = \boldsymbol{x}_0$.
- Then define

$$D_{\widehat{\boldsymbol{F}}} = \{(t, t_0, \boldsymbol{x}_0) \in \mathbb{T} \times \mathbb{T} \times X \mid t \in I_{\widehat{\boldsymbol{F}}}(t_0, \boldsymbol{x}_0)\}$$

  This is the ***domain*** of $\widehat{\boldsymbol{F}}$.
- Then define $\Phi^{\widehat{\boldsymbol{F}}} \colon D_{\widehat{\boldsymbol{F}}} \to X$ by requiring that

$$t \to \Phi^{\widehat{\boldsymbol{F}}}(t, t_0, \boldsymbol{x}_0)$$

  is the solution to $\widehat{\boldsymbol{F}}$ with the initial condition $\boldsymbol{\xi}(t_0) = \boldsymbol{x}_0$.

Ordinary differential equations

- Thus:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^{\widehat{\boldsymbol{F}}}(t, t_0, \boldsymbol{x}_0) = \widehat{\boldsymbol{F}}(t, \Phi^{\widehat{\boldsymbol{F}}}(t, t_0, \boldsymbol{x}_0)).$$

  This is the ***flow*** of $\hat{F}$, and encodes all solutions of the differential equation.

Example

Take $\mathbb{T} = \mathbb{R}$, $X = \mathbb{R}$, $\hat{F}(t, x) = x^2$. This is a separable first-order equation and can easily be solved with the initial state x_0 at time t_0:

$$\xi(t) = \frac{x_0}{1 - x_0(t - t_0)}.$$

Therefore,

$$D_{\hat{F}} = \left\{ (t, t_0, x_0) \;\middle|\; \begin{cases} t \in (-\infty, t_0 + \frac{1}{x_0}), & x_0 > 0, \\ t \in (t_0 + \frac{1}{x_0}, \infty), & x_0 < 0, \\ t \in (-\infty, \infty), & x = 0. \end{cases} \right\}$$

Ordinary differential equations

Example (cont'd)

and

$$\Phi^{\widehat{F}}: D_{\widehat{F}} \to \mathbb{R}$$
$$(t, t_0, x_0) \mapsto \frac{x_0}{1 - x_0(t - t_0)}.$$

- Linear ordinary differential equations will play an important rôle in this course.

Definition

An ***homogeneous linear ordinary differential equation*** is a mapping $\widehat{F}: \mathbb{T} \times \mathsf{X} \to \mathsf{X}$, where $\mathbb{T} \subseteq \mathbb{R}$ is a continuous time-domain, X is an n-dimensional vector space, and $\widehat{F}(t, x) = \mathsf{A}(t)x$, where $\mathsf{A}: \mathbb{T} \to \mathsf{L}(\mathsf{X}; \mathsf{X})$ (linear maps from X to X) is locally integrable.

Ordinary differential equations

- A solution to the homogeneous linear ode is a locally absolutely continuous $\xi: \mathbb{T} \to \mathsf{X}$ satisfying

$$\dot{\xi}(t) = \mathsf{A}(t)\xi(t).$$

- Important facts about homogeneous linear ordinary differential equations:

  1. $D_{\widehat{F}} = \mathbb{T} \times \mathbb{T} \times \mathsf{X}$;
  2. flow is linear in state, i.e., for $(t, t_0) \in \mathbb{T}^2$,

  $$\Phi^{\widehat{F}}(t, t_0, x_0) = \Phi^{\mathsf{c}}_{\mathsf{A}}(t, t_0)x_0$$

  where $\Phi^{\mathsf{c}}_{\mathsf{A}}: \mathbb{T} \times \mathbb{T} \to \mathsf{L}(\mathsf{X}; \mathsf{X})$ is the ***state transition map***.

Reading for Lecture 5

Material related to this lecture can be found in the following sections of the course notes:

1. Sections V-3.2.1 and V-3.2.1.3.
2. Sections V-3.1.3.2, V-3.1.3.3, and V-5.2.1.

Lecture 6
Ordinary differential equations (cont'd)

- Next we consider ordinary differential equations with no dependence on time.

Definition

An ordinary differential equation $\widehat{\boldsymbol{F}}\colon \mathbb{T} \times X \to \mathbb{R}^n$ is ***autonomous*** if there exists $\widehat{\boldsymbol{F}}_0\colon X \to \mathbb{R}^n$ so that $\widehat{\boldsymbol{F}}(t,x) = \widehat{\boldsymbol{F}}_0(x)$.

- The definition just captures you say say that $\widehat{\boldsymbol{F}}$ is independent of time.
- The flow of an autonomous ordinary differential equation has the following property:
$$\Phi^{\widehat{\boldsymbol{F}}}(t, t_0, \boldsymbol{x}_0) = \Psi^{\widehat{\boldsymbol{F}}}(t - t_0, \boldsymbol{x}_0)$$
for some $\Psi^{\widehat{\boldsymbol{F}}}$ defined on some subset of $\mathbb{T} \times X$ to $\mathbb{R}^n$.
- For autonomous equations, one often gives initial conditions at $t_0 = 0$.

Ordinary differential equations

- If $\widehat{F}\colon \mathbb{T} \times \mathsf{X} \to \mathsf{X}$ is an homogeneous linear ordinary differential equation, then it is autonomous if and only if $\widehat{F}(t,x) = \mathsf{A}x$ for $\mathsf{A} \in \mathsf{L}(\mathsf{X};\mathsf{X})$.
- For homogeneous linear ordinary differential equations, autonomous means constant coefficients.
- In this case, we can calculate the state transition map in terms of the operator exponential of A.
- If $\mathsf{L} \in \mathsf{L}(\mathsf{X};\mathsf{X})$, we have

$$\mathsf{e}^{\mathsf{L}} = \mathrm{id}_{\mathsf{X}} + \sum_{n=1}^{\infty} \frac{\mathsf{L}^n}{n!}.$$

- Consider the initial value problem

$$\dot{\xi}(t) = \mathsf{A}\xi(t), \quad \xi(t_0) = x_0.$$

From your past life, you know that the solution is

$$\xi(t) = \mathsf{e}^{\mathsf{A}(t-t_0)}x_0.$$

Ordinary differential equations

- Thus the state transition map is

$$\Phi^{\mathsf{c}}_{\mathsf{A}}(t,t_0) = \mathsf{e}^{\mathsf{A}(t-t_0)}$$

and the flow is

$$\Phi^{\widehat{F}}(t,t_0,x_0) = \mathsf{e}^{\mathsf{A}(t-t_0)} \cdot x_0.$$

- Let us next consider inhomogeneous linear ordinary differential equations.
- In this case, we have

$$\widehat{F}(t,x) = \mathsf{A}(t)x + f(t)$$

for $\mathsf{A} \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{T};\mathsf{L}(\mathsf{X};\mathsf{X}))$ and $f \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{T};\mathsf{X})$.
- The flow is give by the ***variation of constants formula***:

$$\Phi^{\widehat{F}}(t,t_0,x_0) = \Phi^{\mathsf{c}}_{\mathsf{A}}(t,t_0)x_0 + \int_{t_0}^{t} \Phi^{\mathsf{c}}_{\mathsf{A}}(t,\tau)f(\tau)\,\mathrm{d}\tau.$$

- In the constant coefficient case, this becomes

$$\Phi^{\widehat{F}}(t,t_0,x_0) = \mathsf{e}^{\mathsf{A}(t-t_0)}x_0 + \int_{t_0}^{t} \mathsf{e}^{\mathsf{A}(t-\tau)}f(\tau)\,\mathrm{d}\tau.$$

Ordinary difference equations

- We now turn to the discrete-time analogue of ordinary differential equations.

Definition

Let $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ be a discrete time-domain and let $X \subseteq \mathbb{R}^n$ be open. An ***ordinary difference equation*** with time-domain $\mathbb{T}$ and state space X is a mapping $\widehat{\boldsymbol{F}} : \mathbb{T} \times X \to X$. A ***solution*** to the difference equation $\widehat{\boldsymbol{F}}$ is a mapping $\boldsymbol{\xi} : \mathbb{T} \to X$ satisfying

$$\boldsymbol{\xi}(t + \Delta) = \widehat{\boldsymbol{F}}(t, \boldsymbol{\xi}(t)).$$

- The question of existence of solutions for difference equations is resolved by. . . computing them.
- If we have an initial condition $\boldsymbol{\xi}(t_0) = \boldsymbol{x}_0$, then

$$\begin{aligned} \boldsymbol{\xi}(t_0 + \Delta) &= \widehat{\boldsymbol{F}}(t, \boldsymbol{x}_0), \\ \boldsymbol{\xi}(t_0 + 2\Delta) &= \widehat{\boldsymbol{F}}(t, \widehat{\boldsymbol{F}}(t, \boldsymbol{x}_0)), \\ &\vdots \end{aligned}$$

Ordinary difference equations

- Ordinary differential equations have solutions defined for times less than the initial time.
- For an ordinary difference equation with initial condition $\boldsymbol{\xi}(t_0) = \boldsymbol{x}_0$, what is $\boldsymbol{\xi}(t_0 - \Delta)$? It is determined by

$$\boldsymbol{\xi}(t_0) = \widehat{\boldsymbol{F}}(t, \boldsymbol{\xi}(t_0 - \Delta)).$$

  Generally, this expression cannot be solved for $\boldsymbol{\xi}(t_0 - \Delta)$.
- Difference equations are thus meant to "Go forward." If we can go backwards we say the system is ***invertible***.

Example

Take $\mathbb{T} = \mathbb{Z}$, $X = \mathbb{R}$, and $\widehat{F}(t, x) = 0$. Then, for any initial condition $\xi(t_0) = x_0$,

$$\xi(t_0 + k\Delta) = 0 \quad k \in \mathbb{Z}_{>0}.$$

Note that this demonstrates that the lack of invertibility of ordinary difference equations leads to lack of uniqueness of solutions.

Ordinary difference equations

- Although one does not have the same sort of uniqueness as one does for ordinary differential equations, one has "forward uniqueness;" a solution from an initial state $\boldsymbol{x}_0$ at initial time t_0 has a unique solution for all $t \geq t_0$.
- Thus we define the flow

$$\Phi^{\widehat{\boldsymbol{F}}} : \{(t, t_0, \boldsymbol{x}_0) \in \mathbb{T} \times \mathbb{T} \times X \mid t \geq t_0\} \to X$$

  of a difference equations, by just recursively applying the difference equation with the initial condition $\Phi^{\widehat{\boldsymbol{F}}}(t_0, t_0, \boldsymbol{x}_0) = \boldsymbol{x}_0$.
- One may talk about all of the particular sorts of ordinary difference equations as we have done for ordinary differential equations: homogeneous linear, state transition maps, autonomous, homogeneous linear with constant coefficients, and so on. We will not go through this in detail, but will make free use of properties that are entirely analogous to those for ordinary differential equations.

Reading for Lecture 6

Material related to this lecture can be found in the following sections of the course notes:

1. Sections V-3.1.3.1, V-5.2.2, V-5.3.1, and V-5.3.2.
2. Sections V-3.3.3.1, V-3.4.1, V-3.4.1.2, V-3.3.3.2, V-3.3.3.3, V-5.6.1, V-5.6.2, V-5.7.1, and V-5.7.2.

Lecture 7
Distributions

Example

Simple ordinary differential equation:

$$\dot{\eta}(t) = \mu_j(t), \quad \eta(0) = 0,$$

where

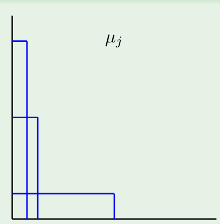
$$\mu_j(t) = \begin{cases} j, & t \in [0, \frac{1}{j}], \\ 0, & \text{otherwise.} \end{cases}$$

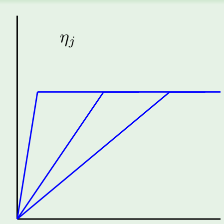
Solve for η_j with input μ_j:

$$\eta_j(t) = \begin{cases} jt, & t \in [0, \frac{1}{j}], \\ 1, & t \in (\frac{1}{j}, \infty). \end{cases}$$

Distributions

Example (cont'd)





As $j \to \infty$,

$$\mu_\infty(t) = \lim_{j\to\infty} \mu_j(t) = \begin{cases} \infty, & t = 0, \\ 0, & \text{otherwise} \end{cases}$$

$$\eta_\infty(t) = \lim_{j\to\infty} \eta_j(t) = \begin{cases} 1, & t \in (0, \infty), \\ 0, & \text{otherwise.} \end{cases}$$

Distributions

Example (cont'd)

Questions:

1. Is η_∞ the output for the input μ_∞?
   No!
2. Is there *any* input μ which gives the output η_∞?
   No! (Not if we restrict to functions as inputs.)

- We will consider a new class of thingies (distributions) which can serve as models for certain physical behaviour.
- Distributions are not functions of time. They are functions of test functions.

Definition

Denote by $\mathscr{D}(\mathbb{R},\mathbb{F})$ the set of infinitely differentiable functions with compact support, i.e., $\mathscr{D}(\mathbb{R};\mathbb{F}) = \mathrm{C}^\infty_{\mathrm{cpt}}(\mathbb{R};\mathbb{F})$. Elements of $\mathscr{D}(\mathbb{R};\mathbb{F})$ are ***test functions***.

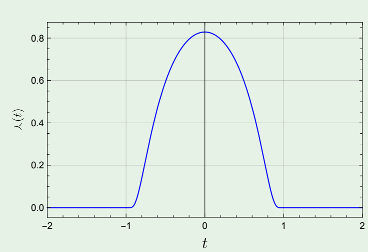
Distributions

Examples

1. $\phi(t) = 0$ is in $\mathscr{D}(\mathbb{R};\mathbb{F})$.
2. Are there nonzero elements of $\mathscr{D}(\mathbb{R};\mathbb{F})$? You will recall from MATH/MTHE 281 functions like the following:

$$⅄(t) = \begin{cases} \mathrm{e}^{-1/(1-t^2)}, & t \in (-1,1), \\ 0, & \text{otherwise} \end{cases}$$

   whose graph is



Distributions

Examples (cont'd)

This function clearly has compact support and is infinitely differentiable, except possibly at $t = \pm 1$. One calculates

$$\frac{\mathrm{d}}{\mathrm{d}t} \curlywedge(t) = \frac{1}{P(t)} \curlywedge(t),$$

where P is a polynomial that vanishes at $t = \pm 1$. As exponentials go to infinity fast than polynomials, the derivative vanishes at $t = \pm 1$. This process can be continued to show that the derivatives of $\curlywedge$ are defined and equal to zero at $t = \pm 1$. Thus $\curlywedge$ is infinitely differentiable.

- Test functions themselves are not of specific interest; they are "cannon fodder" in some sense, since we will primarily think of them as being arguments for the things we actually care about.

Distributions

- But we do need a notion of convergence of a sequence of test functions.

Definition

A sequence $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$ if

(i) there exists $K \subseteq \mathbb{R}$ compact such that $\mathrm{supp}(\phi_j) \subseteq K$ for $j \in \mathbb{Z}_{>0}$;

(ii) $(\phi_j^{(k)})_{j\in\mathbb{Z}_{>0}}$ converges uniformly to zero for every $k \in \mathbb{Z}_{\geq 0}$, i.e., ϕ_j and all of its derivatives converge uniformly to zero.

- If you like norms, this sort of convergence is not norm convergence.
- If you like metrics, this sort of convergence is not metric convergence.
- If you like strict inductive limits of locally convex spaces. . . then you are in luck!

Reading for Lecture 7

Material related to this lecture can be found in the following sections of the course notes:

1. Section IV-3.1.
2. Section IV-3.2.1.

Lecture 8
Distributions (cont'd)

Definition

A ***distribution*** is a mapping $\theta\colon \mathscr{D}(\mathbb{R};\mathbb{F}) \to \mathbb{F}$ such that

(i) it is linear: i.e.,

$$\begin{aligned}\theta(\phi_1+\phi_2) &= \theta(\phi_1)+\theta(\phi_2),\\ \theta(a\phi) &= a(\theta(\phi).\end{aligned}$$

(ii) it is continuous: i.e., if $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$, then $(\theta(\phi_j))_{j\in\mathbb{Z}_{>0}}$ converges to zero.

Examples

1. If $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$, then define

$$\begin{aligned}\theta_f\colon \mathscr{D}(\mathbb{R};\mathbb{F}) &\to \mathbb{F}\\ \phi &\mapsto \int_{\mathbb{R}} f(t)\phi(t)\,\mathrm{d}t.\end{aligned}$$

Distributions

Examples (cont'd)

Claim that θ_f is a distribution.
Step 1 θ_f is well-defined, i.e., the integral exists.
Proof:

$$\begin{aligned}\int_{\mathbb{R}} |f(t)\phi(t)|\,\mathrm{d}t &= \int_{\mathrm{supp}(\phi)} |f(t)||\phi(t)|\,\mathrm{d}t\\ &= \|\phi\|_\infty \int_{\mathrm{supp}(\phi)} |f(t)|\,\mathrm{d}t\\ &< \infty\end{aligned}$$

because $f \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{R};\mathbb{F})$.
Step 2 θ_f is linear.
Proof: Linearity of the integral.

Distributions

Examples (cont'd)

Step 3 θ_f is continuous.
Proof: Let $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ converge to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$. Let $K \subseteq \mathbb{R}$ be compact and such that $\mathrm{supp}(\phi_j) \subseteq K$, $j \in \mathbb{Z}_{>0}$. Then

$$\begin{aligned}|\theta_f(\phi_j)| = \left|\int_{\mathbb{R}} f(t)\phi_j(t)\,\mathrm{d}t\right| &\leq \int_{\mathbb{R}} |f(t)\phi_j(t)|\,\mathrm{d}t\\ &= \int_K |f(t)\phi_j(t)|\,\mathrm{d}t \leq \|\phi_j\|_\infty \int_K |f(t)|\,\mathrm{d}t\\ &\leq M\|\phi_j\|_\infty.\end{aligned}$$

Thus,

$$\lim_{j\to\infty} |\theta(\phi_j)| \leq \lim_{j\to\infty} M\|\phi_j\|_\infty = 0.$$

Thus, every locally integrable function f defines a distribution θ_f.
NB. The map $f \mapsto \theta_f$ is injective.

Distributions

Examples (cont'd)

2. Define $\delta\colon \mathscr{D}(\mathbb{R};\mathbb{F}) \to \mathbb{F}$ by $\delta(\phi) = \phi(0)$. Claim that δ is a distribution.
   Step 1 δ is linear.
   Proof: Obvious.
   Step 2 δ is continuous.
   Proof: Let $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence converging to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$. Then

$$\begin{aligned} & |\delta(\phi_j)| = |\phi_j(0)| \\ \implies & \lim_{j\to\infty} |\delta(\phi_j)| = \lim_{j\to\infty} |\phi_j(0)| = 0. \end{aligned}$$

   We call δ the ***Dirac δ-distribution***. Can also define $\delta_{t_0}\colon \mathscr{D}(\mathbb{R};\mathbb{F}) \to \mathbb{F}$ by $\delta_{t_0}(\phi) = \phi(t)$.
3. Let θ be a distribution. The ***derivative*** of θ is the distribution θ' or $\theta^{(1)}$ defined by

$$\theta^{(1)}(\phi) = -\theta(\phi^{(1)}).$$

Distributions

Examples (cont'd)

We can define higher-derivatives recursively

$$\theta^{(k)}(\phi) = (-1)^k \theta(\phi^{(k)}).$$

Claim that $\theta^{(1)}$ is a distribution.
Step 1 $\theta^{(1)}$ is well-defined.
Proof: Must check that $\phi^{(1)} \in \mathscr{D}(\mathbb{R};\mathbb{F})$. This is clear.
Step 2 $\theta^{(1)}$ is linear.
Proof:

$$\begin{aligned} \theta^{(1)}(\phi_1 + \phi_2) &= -\theta((\varphi_1 + \varphi_2)^{(1)}) \\ &= -\theta(\phi_1^{(1)} + \phi_2^{(1)}) \\ &= \theta^{(1)}(\phi_1) + \theta^{(1)}(\phi_2) \\ \theta(a\phi) &= a\theta^{(1)}(\phi). \end{aligned}$$

Distributions

Examples (cont'd)

Step 3 $\theta^{(1)}$ is continuous.
Proof: Let $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ be a sequence converging to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$. Then

$$|\theta^{(1)}(\phi_j)| = |\theta(\phi_j^{(1)})|.$$

Because $(\phi_j^{(1)})_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$ and because θ is continuous,

$$\lim_{j\to\infty}|\theta^{(1)}(\phi_j)| = \lim_{j\to\infty}|\theta(\phi^{(1)})| = 0.$$

Why is the definition of the derivative of a distribution as it is? Let $f \in \mathrm{C}^1(\mathbb{R};\mathbb{F})$. We claim that $\theta_f^{(1)} = \theta_{f^{(1)}}$. Indeed let $\phi \in \mathscr{D}(\mathbb{R};\mathbb{F})$ and

Distributions

Examples (cont'd)

compute

$$\begin{aligned}\theta_{f^{(1)}}(\phi) &= \int_{\mathbb{R}} f^{(1)}(t)\phi(t)\,\mathrm{d}t\\ &= f(t)\phi(t)\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} f(t)\phi^{(1)}(t)\,\mathrm{d}t\\ &= -\int_{\mathbb{R}} f(t)\phi^{(1)}(t)\,\mathrm{d}t\\ &= -\theta_f(\phi^{(1)}) = \theta_f^{(1)}(\phi).\end{aligned}$$

4. Generally, one cannot multiply distributions as one multiplies functions. However, one can multiply a distribution θ by a function $f \in \mathrm{C}^\infty(\mathbb{R};\mathbb{F})$ as follows:

$$(f\theta)(\phi) = \theta(f\phi).$$

Distributions

Examples (cont'd)

To show that $f\theta$ is a distribution, one must show (a) that it is well-defined (meaning we must show that $f\phi \in \mathscr{D}(\mathbb{R};\mathbb{F})$, which is clear, (b) that it is linear (this is obvious), and (c) it is continuous. Continuity can be proved by appeal to the higher-order Leibniz Rule:

$$(fg)^{(k)} = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} f^{(j)} g^{(k-j)}.$$

We leave the details to the reader.

Reading for Lecture 8

Material related to this lecture can be found in the following sections of the course notes:

1. Sections IV-3.2.2, IV-3.2.3, and IV-3.2.6.

Lecture 9
Distributions (cont'd)

- We showed that, if $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$, then θ_f is a distribution. Looking at the proof of continuity of θ_f, we used, in an essential way, the fact that, if $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$, then there exists a compact $K \subseteq \mathbb{R}$ such that $\mathrm{supp}(\phi_j) \subseteq K$. Thus this condition on convergence to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$ is necessary for θ_f to be a distribution.
- We also defined, for a distribution θ, its kth derivative. The proof of continuity of $\theta^{(k)}$ used, in an essential way, the fact that, if $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$, then so does $(\phi_j^{(k)})_{j\in\mathbb{Z}_{>0}}$.
- *Punchline:* The definition of convergence to zero in $\mathscr{D}(\mathbb{R};\mathbb{F})$ is forced upon us by requiring that
  1. θ_f is a distribution for every $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$ and
  2. all distributions be differentiable.

Distributions

- We denote by $\mathscr{D}'(\mathbb{R};\mathbb{F})$ the set of distributions.
- Note that $\mathscr{D}'(\mathbb{R};\mathbb{F})$ is a subset of the set of linear mappings from $\mathscr{D}(\mathbb{R};\mathbb{F})$ to $\mathbb{F}$. This set of *all* linear mappings is called the ***algebraic dual*** of $\mathscr{D}(\mathbb{R};\mathbb{F})$.
- Then $\mathscr{D}'(\mathbb{R};\mathbb{F})$ is the ***topological dual***.
- We can talk about convergence of sequences of distributions.

Definition

A sequence $(\theta_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{D}'(\mathbb{R};\mathbb{F})$ ***converges*** to $\theta \in \mathscr{D}'(\mathbb{R};\mathbb{F})$ if $(\theta_j(\phi))_{j\in\mathbb{Z}_{>0}}$ converges to $\theta(\phi)$ for all $\phi \in \mathscr{D}(\mathbb{R};\mathbb{F})$.

Distributions

Example

Define $f_j : \mathbb{R} \to \mathbb{F}$ by

$$f_j(t) = \begin{cases} j, & t \in [0, \frac{1}{j}], \\ 0, & \text{otherwise}. \end{cases}$$

Claim that $(\theta_{f_j})_{j \in \mathbb{Z}_{>0}}$ converges to δ. Indeed, let $\phi \in \mathscr{D}(\mathbb{R}; \mathbb{F})$. Define

$$\epsilon_j = \sup\{|\phi(t) - \phi(0)| \mid t \in [0, \tfrac{1}{j}]\}.$$

Thus

$$\phi(0) - \epsilon_j \leq \phi(t) \leq \phi(0) + \epsilon_j, \qquad t \in [0, \tfrac{1}{j}],$$

and integration $j \int_0^{1/j} \mathrm{d}t$ gives

$$\phi(0) - \epsilon_j \leq \theta_{f_j}(\phi) \leq \phi(0) + \epsilon_j.$$

As $j \to \infty$, $\epsilon_j \to 0$, and so $\lim_{j\to\infty} \theta_{f_j}(\phi) = \phi(0) = \delta(\phi)$, as claimed.

Distributions

Example

Let

$$1_{\geq 0}(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0. \end{cases}$$

Claim that $\theta^{(1)}_{1_{\geq 0}} = \delta$. For $\phi \in \mathscr{D}(\mathbb{R}; \mathbb{F})$,

$$\begin{aligned} \theta^{(1)}_{1_{\geq 0}}(\phi) &= -\theta_{1_{\geq 0}}(\phi^{(1)}) \\ &= -\int_{\mathbb{R}} 1_{\geq 0}(t)\phi^{(1)}(t)\,\mathrm{d}t = -\int_0^\infty \phi^{(1)}(t)\,\mathrm{d}t \\ &= -\phi(t)\Big|_0^\infty = \phi(0) = \delta(\phi), \end{aligned}$$

as claimed.

Reading for Lecture 9

Material related to this lecture can be found in the following sections of the course notes:

1. Section IV-3.2.5.

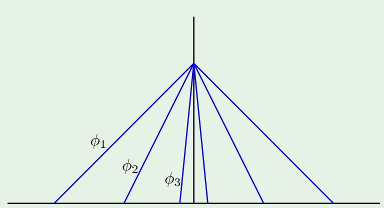
Lecture 10
Distributions (cont'd)

Example

We claim that there is no $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ such that

$$\int_{\mathbb{R}} f(t)\phi(t) = \phi(0), \qquad \phi \in \mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R},\mathbb{R}).$$

Suppose there is such an f. Define $\phi_j(t) \in \mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathbb{R})$, $j \in \mathbb{Z}_{>0}$, as depicted here:



Distributions

Example

Suppose that

$$\int_{\mathbb{R}} f(t)\phi_j(t)\,\mathrm{d}t = \phi_j(0) = 1.$$

By the Dominated Convergence Theorem,

$$\lim_{j\to\infty}\int_{\mathbb{R}} f(t)\phi_j(t)\,\mathrm{d}t = \int_{\mathbb{R}} f(t)\lim_{j\to\infty}\phi_j(t)\,\mathrm{d}t = 0.$$

But we also have, by assumption,

$$\lim_{j\to\infty}\int_{\mathbb{R}} f(t)\phi_j(t)\,\mathrm{d}t = \lim_{j\to\infty}\phi_j(0) = 1.$$

Contradiction shows that no such f exists.
This can easily be adapted to $\phi_j \in \mathsf{C}^\infty_{\mathsf{cpt}}(\mathbb{R};\mathbb{R})$. Therefore, $\delta \neq \theta_f$ for any $f \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{R};\mathbb{R})$.

Distributions and ordinary differential equations

- Distributions are well suited to the study of linear differential equations (ordinary and partial) with constant coefficients.
- Sometimes, if one can prove the existence of a solution that is a distribution, this can be used to prove the existence of solutions in the usual sense.
- We first consider the scalar case. We represent a kth-order scalar linear ordinary differential equation with constant coefficients by

$$\widehat{F}(t, x, x^{(1)}, x^{(2)}, \dots, x^{(k-1)}) = -a_0 x - a_1 x^{(1)} - a_2 x^{(2)} - \cdots - a_{k-1}x^{(k-1)}, \quad x^{(j)} \in \mathbb{R}.$$

- A solution is $\xi\colon \mathbb{T} \to \mathbb{R}$ satisfying

$$\frac{\mathrm{d}^k\xi}{\mathrm{d}t^k}(t) = \widehat{F}\left(t, \xi(t), \frac{\mathrm{d}\xi}{\mathrm{d}t}(t), \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2}(t), \dots, \frac{\mathrm{d}^{k-1}\xi}{\mathrm{d}t^{k-1}}(t)\right)$$
$$\implies \frac{\mathrm{d}^k\xi}{\mathrm{d}t^k}(t) + a_{k-1}\frac{\mathrm{d}^{k-1}\xi}{\mathrm{d}t^{k-1}}(t) + \cdots + a_1\frac{\mathrm{d}\xi}{\mathrm{d}t}(t) + a_0\xi(t) = 0.$$

Distributions and ordinary differential equations

- We consider equations like this, but with distributions on the right-hand side, i.e., an inhomogeneous equation with the inhomogeneous term being a distribution. We will ask for solutions that are themselves distributions.
- If the solution is $\theta \in \mathscr{D}'(\mathbb{R};\mathbb{R})$, then

$$\theta^{(k)} + a_{k-1}\theta^{(k-1)} + \cdots + a_1\theta^{(1)} + a_0\theta = \beta \tag{1}$$

  for $\beta \in \mathscr{D}'(\mathbb{R};\mathbb{R})$.
- To study this equation initial conditions are not meaningful. Instead we consider restrictions on the "support" of solutions.

Definition

Let $\theta \in \mathscr{D}'(\mathbb{R};\mathbb{R})$ and let $I \subseteq \mathbb{R}$ be an open interval. Say that θ ***vanishes on I*** if $\theta(\phi) = 0$ for all $\phi \in \mathscr{D}(\mathbb{R};\mathbb{R})$ with $\mathrm{supp}(\phi) \subseteq I$. The ***support*** of θ is

$$\mathrm{supp}(\theta) = \mathbb{R} \setminus \cup\{I \subseteq \mathbb{R} \mid \theta \text{ vanishes on } I\}.$$

Distributions and ordinary differential equations

Examples

1. $\mathrm{supp}(\theta_f) = \mathrm{supp}(f), f \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{R};\mathbb{R})$.
2. $\mathrm{supp}(\delta) = \{0\}$. Indeed, since $\delta(\phi) = \phi(0)$,

$$\begin{aligned} \delta(\phi) = 0 \quad &\Longleftrightarrow \quad \mathrm{supp}(\phi) \subseteq (-\infty,0) \cup (0,\infty) \\ \Longrightarrow \quad \mathrm{supp}(\delta) = \{0\}. \end{aligned}$$

- Denote by

$$\begin{aligned} \mathscr{D}'_+(\mathbb{R};\mathbb{R}) &= \{\theta \in \mathscr{D}'(\mathbb{R};\mathbb{R}) \mid \inf\mathrm{supp}(\theta) > -\infty\}, \\ \mathscr{D}'_-(\mathbb{R};\mathbb{R}) &= \{\theta \in \mathscr{D}'(\mathbb{R};\mathbb{R}) \mid \sup\mathrm{supp}(\theta) < \infty\}. \end{aligned}$$

Distributions and ordinary differential equations

Theorem

Let $\widehat{F}$ be a kth-order scalar linear ordinary differential equation with constant coefficients and let $\beta \in \mathscr{D}'(\mathbb{R};\mathbb{R})$. Denote by $\mathrm{Sol}(\widehat{F},\beta)$ the set of solutions of (1). Then:

(i) $\mathrm{card}(\mathrm{Sol}(\widehat{F},\beta)) \geq 2$*;*

(ii) if $\beta \in \mathscr{D}'_+(\mathbb{R};\mathbb{R})$*, then*

$$\mathrm{card}(\mathrm{Sol}(\widehat{F},\beta) \cap \mathscr{D}'_+(\mathbb{R};\mathbb{R})) = 1;$$

(iii) if $\beta \in \mathscr{D}'_-(\mathbb{R};\mathbb{R})$*, then*

$$\mathrm{card}(\mathrm{Sol}(\widehat{F},\beta) \cap \mathscr{D}'_-(\mathbb{R};\mathbb{R})) = 1.$$

- We will consider parts (ii) and (iii) of the theorem in the case $\beta = \delta$.
- It turns out that all other β's follow from this, but this involves delving into things a little outside the scope of what we do here.

Reading for Lecture 10

Material related to this lecture can be found in the following sections of the course notes:

1. Proposition IV-3.1.1 and Section IV-3.2.4.
2. Section V-4.4.3.

Lecture 11
Distributions and ordinary differential equations (cont'd)

- To determine the unique solution in $\mathscr{D}'_+(\mathbb{R};\mathbb{R})$ to (1) with $\beta=\delta$, let $\xi\colon\mathbb{R}\to\mathbb{R}$ be the unique solution to the initial value problem:

$$\frac{\mathrm{d}^k\xi}{\mathrm{d}t^k}(t)+a_{k-1}\frac{\mathrm{d}^{k-1}\xi}{\mathrm{d}t^{k-1}}(t)+\cdots+a_1\frac{\mathrm{d}\xi}{\mathrm{d}t}(t)+a_0\xi(t)=0,$$
$$\xi(0)=\frac{\mathrm{d}\xi}{\mathrm{d}t}(0)=\cdots=\frac{\mathrm{d}^{k-2}\xi}{\mathrm{d}t^{k-2}}(0)=0,\ \frac{\mathrm{d}^{k-1}\xi}{\mathrm{d}t^{k-1}}(0)=1.$$

- Denote by

$$1_{\geq 0}(t)=\begin{cases}1, & t\geq 0,\\ 0, & t<0\end{cases}$$

the unit step function.
- We will show that $\theta_{1_{\geq 0}\xi}$ is the unique solution in $\mathrm{D}'_+(\mathbb{R};\mathbb{R})$ to (1) with $\beta=\delta$.

Distributions and ordinary differential equations

- First note that $\theta_{1_{\geq 0}\xi}=\xi\theta_{1_{\geq 0}}$. Indeed, for $\phi\in\mathscr{D}(\mathbb{R};\mathbb{R})$,

$$\theta_{1_{\geq 0}\xi}(\phi)=\int_{\mathbb{R}}1_{\geq 0}(t)\xi(t)\phi(t)\,\mathrm{d}t$$
$$=\theta_{1_{\geq 0}}(\xi\phi)=\xi\theta_{1_{\geq 0}}(\phi).$$

Fact

If $\theta\in\mathscr{D}'(\mathbb{R};\mathbb{R})$ and if $f\in\mathrm{C}^\infty(\mathbb{R};\mathbb{R})$, then

$$(f\theta)^{(1)}=f^{(1)}\theta+f\theta^{(1)}.$$

Proof.

Note that, for $\phi\in\mathscr{D}(\mathbb{R};\mathbb{R})$, $(f\phi)^{(1)}=f^{(1)}\phi+f\phi^{(1)}$. Thus

$$(f\theta)^{(1)}(\phi)=-f\theta(\phi^{(1)})=-\theta(f\phi^{(1)})=-\theta((f\phi)^{(1)})+\theta(f^{(1)}\phi)$$
$$=\theta^{(1)}(f\phi)+f^{(1)}\theta(\phi)=f\theta^{(1)}(\phi)+f^{(1)}\theta(\phi)=(f\theta^{(1)}+f^{(1)}\theta)(\phi).$$

□

Distributions and ordinary differential equations

- Then we calculate:

$$(\xi\theta_{1_{\geq 0}})^{(1)} = \xi^{(1)}\theta_{1_{\geq 0}} + \xi\theta_{1_{\geq 0}}^{(1)}$$
$$= \xi^{(1)}\theta_{1_{\geq 0}} + \xi\delta.$$

- But

$$(\xi\delta)(\phi) = \delta(\xi\varphi) = \xi(0)\phi(0) = 0.$$

- Thus $(\xi\theta_{1_{\geq 0}})^{(1)} = \xi^{(1)}\theta_{1_{\geq 0}}$.
- Then we recursively have

$$(\xi\theta_{1_{\geq 0}})^{(j)} = \xi^{(j)}\theta_{1_{\geq 0}}, \qquad j \in \{0, 1, \dots, k-2\}.$$

- Then

$$(\xi\theta_{1_{\geq 0}})^{(k)} = \xi^{(k)}\theta_{1_{\geq 0}} + \xi^{(k-1)}\theta_{1_{\geq 0}}^{(1)}$$
$$= \xi^{(k)}\theta_{1_{\geq 0}} + \delta.$$

Distributions and ordinary differential equations

- We now consider the problem of systems of linear ordinary differential equations with constant coefficients and with a distribution for an inhomogeneous term.
- The homogeneous equation is thus $\widehat{F}(t,x) = \mathsf{A}x$, $x \in \mathsf{X}$ (a finite-dimensional $\mathbb{R}$-vector space).
- We look for distributional solutions to

$$\theta^{(1)} = \mathsf{A}(\theta) + \beta. \qquad (2)$$

- We must make sense of the symbols in this equation before we do anything else.

Definition

Let X be a finite-dimensional $\mathbb{R}$-vector space. A ***X-valued distribution*** is a continuous linear mapping $\theta\colon \mathscr{D}(\mathbb{R};\mathbb{R}) \to \mathsf{X}$. Denote the set of these distributions by $\mathscr{D}'(\mathbb{R};\mathsf{X})$.

Distributions and ordinary differential equations

- We shall now give a list of constructions that are necessary to present solutions to (2).

Constructions

1. Let $x_0 \in \mathsf{X}$ and let $\theta \in \mathscr{D}'(\mathbb{R};\mathbb{R})$. Define $x_0 \otimes \theta \in \mathscr{D}'(\mathbb{R};\mathsf{X})$ by
$$x_0 \otimes \theta(\phi) = \theta(\phi)x_0.$$
The symbol "$\otimes$" is pronounced "tensor" (not "tenser") and the symbol $x_0 \otimes \theta$ should be thought of as the product of x_0 and θ.
2. Let $\mathsf{A} \in \mathsf{L}(\mathsf{X};\mathsf{X})$ and $\theta \in \mathscr{D}'(\mathbb{R};\mathsf{X})$. Define $\mathsf{A}(\theta) \in \mathscr{D}'(\mathbb{R};\mathsf{X})$ by
$$\mathsf{A}(\theta)(\phi) = \mathsf{A}(\theta(\phi)).$$

Reading for Lecture 11

Material related to this lecture can be found in the following sections of the course notes:

1. Section V-4.4.3.
2. Sections IV-3.2.12 and V-5.3.3.2.

Lecture 12
Distributions and ordinary differential equations (cont'd)

Constructions (cont'd)

3. If $\xi \in \mathrm{C}^\infty(\mathbb{R};\mathsf{X})$, then we can choose a basis $(e_1, \dots, e_n)$ for X and write

$$\xi(t) = \xi_1(t)e_1 + \cdots + \xi_n(t)e_n.$$

for $\xi_1, \dots, \xi_n \in \mathrm{C}^\infty(\mathbb{R};\mathbb{R})$. If $\theta \in \mathscr{D}'(\mathbb{R};\mathbb{R})$, define

$$\xi \otimes \theta \in \mathscr{D}'(\mathbb{R};\mathsf{X}).$$

by

$$(\xi \otimes \theta)(\phi) = (\xi_1\theta)(\phi) + \cdots + (\xi_n\theta)(\phi).$$

This is the vector version of multiplying a distribution by a C^∞-function.

Distributions and ordinary differential equations

- Then, taking $\xi(t) = \mathrm{e}^{\mathsf{A}t}x_0$, we claim that $\theta_{1_{\geq 0}\xi}$ is the unique solution in $\mathscr{D}'_+(\mathbb{R};\mathsf{X})$ to (2) with $\beta = x_0 \otimes \delta$.
- Indeed, we have $\theta_{1_{\geq 0}\xi} = \xi \otimes \theta_{1_{\geq 0}}$ (this is easily and directly verified), and so we compute

$$\begin{aligned}(\xi \otimes \theta_{1_{\geq 0}})^{(1)} &= \xi^{(1)} \otimes \theta_{1_{\geq 0}} + \xi \otimes \theta_{1_{\geq 0}}^{(1)} \\ &= (\mathsf{A}\xi) \otimes \theta_{1_{\geq 0}} + \xi \otimes \delta.\end{aligned}$$

- Now we can easily and directly verify:
  1. $(\mathsf{A}\xi) \otimes \theta_{1_{\geq 0}} = \mathsf{A}(\xi \otimes \theta_{1_{\geq 0}})$;
  2. $\xi \otimes \delta = \xi(0) \otimes \delta = x_0 \otimes \delta$.
- Thus

$$(\xi \otimes \theta_{1_{\geq 0}})^{(1)} = \mathsf{A}(\xi \otimes \theta_{1_{\geq 0}}) + x_0 \otimes \delta,$$

as claimed.
- Thus $\xi \otimes \theta_{1_{\geq 0}}$ is the result of doing nothing up to time 0, applying an impulse in the direction of x_0 at time 0, and then doing nothing thereafter.

Convolution

Convolution $f * g$

Continuous-time	Discrete-time
$f * g(t) = \int_{\mathbb{R}} f(t-s)g(s)\,\mathrm{d}s$	$f * g(k\Delta) = \sum_{j=-\infty}^{\infty} f((k-j)\Delta)g(j\Delta)$

-
- We will start with the continuous-time case.
- Denote
$$D(f,g) = \{t \in \mathbb{R} \mid\ s \mapsto f(t-s)g(s) \text{ is in } \mathsf{L}^1(\mathbb{R};\mathbb{F})\}.$$
- Say that (f,g) is ***convolvable*** if $\mathbb{R} \setminus D(f,g)$ has measure zero.
- Exact conditions under which (f,g) is convolvable are not really meaningful. Instead, one hopes to give conditions on f and g which ensure that (f,g) is convolvable, and which give some properties of $f * g$.

Convolution

- Before we get to this, we consider an example illustrating how convolution works.

Example

Consider:
$$f(t) = \begin{cases} 1, & t \in [-1,1], \\ 0, & \text{otherwise}. \end{cases}$$
Calculate $f * f$. First we note that
$$\begin{aligned} f(t-s) &= \begin{cases} 1, & t-s \in [-1,1], \\ 0, & \text{otherwise}, \end{cases} \\ &= \begin{cases} 1, & -s \in [-t-1,-t+1], \end{cases} \\ &= \begin{cases} 1, & s \in [t-1,t+1]. \end{cases} \end{aligned}$$

Convolution

Example (cont'd)

Thus

$$\begin{aligned} f * f(t) &= \int_{\mathbb{R}} f(t-s) f(s)\,\mathrm{d}s \\ &= \int_{-1}^{1} f(t-s)\,\mathrm{d}s \\ &= \lambda([-1,1] \cap [t-1, t+1]) \\ &= \begin{cases} 2 - |t|, & t \in [-2,2], \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Observation: Convolution "smears support" and "smooths" the functions.

Reading for Lecture 12

Material related to this lecture can be found in the following sections of the course notes:

1. Sections IV-3.2.12 and V-5.3.3.2.
2. Section IV-4.1.1.

Lecture 13
Convolution (cont'd)

- Properties of continuous-time convolution

$$f * g(t) = \int_{\mathbb{R}} f(t-s)g(s)\,\mathrm{d}s.$$

1. If (f,g) is convolvable, then (g,f) is convolvable, and $f*g = g*f$.
2. If (f,g) and (f,h) are convolvable, then $(f,g+h)$ is convolvable and $f*(g+h) = f*g + f*h$.
3. It is not generally true that

$$(f*g)*h = f*(g*h).$$

4. If (f,g) is convolvable and if $f,g \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R},\mathbb{F})$, then $f*g \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$.
5. For $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$, denote by

$$\sigma(f) = \sup\{t \in \mathbb{R} \mid f(s) = 0 \text{ for almost every } s \leq t\}.$$

If $f,g \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$, if $\sigma(f),\sigma(g) > -\infty$, and if (f,g) is convolvable, then

$$f*g(t) = \begin{cases} \int_{\sigma(g)}^{t-\sigma(f)} f(t-s)g(s)\,\mathrm{d}s, & t \geq \sigma(f)+\sigma(g), \\ 0, & \text{otherwise.} \end{cases}$$

Convolution

Indeed, suppose that $t < \sigma(f)+\sigma(g)$. First consider $s \leq \sigma(g)$ in which case $f(t-s)g(s) = 0$. In the other case, with $s \geq \sigma(g)$, we have

$$t - s < \sigma(f) + \sigma(g) - s \leq \sigma(f) \quad \Longrightarrow \quad f(t-s)g(s) = 0.$$

If $t \geq \sigma(f)+\sigma(g)$, consider

$$s > t - \sigma(f) \quad \Longrightarrow \quad t - s < \sigma(f) \quad \Longrightarrow \quad f(t-s)g(s) = 0.$$

6. If $\sigma(f),\sigma(g) \geq 0$, this simplifies to

$$f*g(t) = \begin{cases} \int_0^t f(t-s)g(s)\,\mathrm{d}s, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Compare this to the following term from the variation of constants formula:

$$\int_0^t \mathrm{e}^{\mathsf{A}(t-\tau)} f(\tau)\,\mathrm{d}\tau.$$

We can see that this expression is, in fact, a convolution if one considers all signals as having their values set to zero for negative time.

Convolution

- Let us now turn to the case of when a pair of signals (f, g) is convolvable. There are no useful general conditions, so we can only give special cases.

Theorem

If $f, g \in \mathsf{L}^1(\mathbb{R};\mathbb{F})$, then

(i) (f, g) is convolvable,

(ii) $f * g \in \mathsf{L}^1(\mathbb{R};\mathbb{F})$, and

(iii) $(f * g) * h = f * (g * h)$.

- This shows that $\mathsf{L}^1(\mathbb{R};\mathbb{F})$ is a commutative ring with the convolution product. It has some not so great properties as a ring, however. Here are some additional properties:
  1. There is no unit, i.e., there is no signal $u \in \mathsf{L}^1(\mathbb{R};\mathbb{F})$ such that $u * f = f$ for every $f \in \mathsf{L}^1(\mathbb{R};\mathbb{F})$. If there were a unit, what property should it have?

$$f * u(t) = \int_{\mathbb{R}} f(t-s)u(s)\,\mathrm{d}s = f(t).$$

Convolution

Take $t = 0$. Then

$$\int_{\mathbb{R}} f(-s)u(s)\,\mathrm{d}s = f(0).$$

This looks quite like the behaviour of the Dirac δ-function. But there is no such "function."

2. There are nonzero $f, g \in \mathsf{L}^1(\mathbb{R},\mathbb{F})$ such that $f * g = 0$. (This means that $\mathsf{L}^1(\mathbb{R};\mathbb{F})$ is not an "integral domain.")
3. The convolution product is "surjective," i.e., if $h \in \mathsf{L}^1(\mathbb{R};\mathbb{F})$, then there exists $f, g \in \mathsf{L}^1(\mathbb{R};\mathbb{F})$ such that $f * g = h$.

Convolution

Theorem

If $f, g, h \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$, then

(i) (f, g) is convolvable,

(ii) $(f * g) \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$, and

(iii) $(f * g) * h = f * (g * h)$.

- The ring properties of $\mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$ are the same as those of $\mathsf{L}^1(\mathbb{R}; \mathbb{F})$, except for 2, which is replaced with
  - 2' $\mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$ is an integral domain, i.e., if $f * g = 0$, then either $f = 0$ or $g = 0$. This follows from the Titschmarch Convolution Theorem (very hard to prove).

Convolution

- We now consider pairs of convolvable signals where this is not necessarily symmetry in the properties of each signals in the pair, and also not necessarily symmetry with the properties of the convolution.

Theorem

Let $p, q, r \in [1, \infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. If $f \in \mathsf{L}^p(\mathbb{R}; \mathbb{F})$ and if $g \in \mathsf{L}^q(\mathbb{R}; \mathbb{F})$, then (f, g) is convolvable and $f * g \in \mathsf{L}^r(\mathbb{R}; \mathbb{F})$.

- This follows from "Young's inequality" whose precise form we shall state in the next lecture.
- Let's look at some special cases. If $p \in [1, \infty]$, let $p' \in [1, \infty]$ be defined by $\frac{1}{p} + \frac{1}{p'} = 1$. Call p' the ***conjugate index*** of p.
  1. $p = p$ and $q = p'$: Then $r = \infty$. In this case, more is true, namely that $f * g \in \mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R}; \mathbb{F})$ if $f \in \mathsf{L}^p(\mathbb{R}; \mathbb{F})$ and $g \in \mathsf{L}^{p'}(\mathbb{R}; \mathbb{F})$.
  2. $p = p$ and $q = 1$: Then $r = p$.

Reading for Lecture 13

Material related to this lecture can be found in the following sections of the course notes:

1. Sections IV-4.1.2, IV-4.2.1, IV-4.2.3, and IV-4.2.2.

Lecture 14
Convolution (cont'd)

- Our last result concerning convolvable pairs of continuous-time signals is a natural extension of our previous result to the case of signals with support in $\mathbb{R}_{\geq 0}$.

Theorem

Let $p, q, r \in [1, \infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. If $f \in \mathsf{L}^p_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$ and if $g \in \mathsf{L}^q_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$, then (f, g) is convolvable and $f * g \in \mathsf{L}^r_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$.

- We have the same special cases as previously.
    1. $p = p$ and $q = p'$: Then $r = \infty$. In this case, more is true, namely that $f * g \in \mathsf{C}^0(\mathbb{R}_{\geq 0}; \mathbb{F})$ if $f \in \mathsf{L}^p(\mathbb{R}_{\geq 0}; \mathbb{F})$ and $g \in \mathsf{L}^{p'}(\mathbb{R}_{\geq 0}; \mathbb{F})$.
    2. $p = p$ and $q = 1$: Then $r = p$.

Convolution

- Now we consider the continuity of convolution. There are (at least) two sorts of continuity one can consider.
  1. If we fix a signal g, we can consider the continuity of the mapping $f \mapsto f * g$. In this case, the continuity is of a type we are familiar with as f will be a member of some signal space in which we understand convergence (i.e., a normed space or a space whose topology is defined by seminorms), as will be $f * g$. In this case, continuity amounts to: if $(f_j)_{j\in\mathbb{Z}_{>0}}$ converges to zero, then $(f_j * g)_{j\in\mathbb{Z}_{>0}}$ converges to zero.
  2. We can also consider continuity of $(f, g) \mapsto f * g$. We shall not give here a precise definition of what continuity of a mapping like this means; the mapping is not linear, but rather is bilinear, i.e., linear in each entry. All we shall say is:
     1. if the mapping $(f, g) \mapsto f * g$ is continuous, then, if we fix g, the mapping $f \mapsto f * g$ is also continuous;
     2. we shall see that all of our spaces of convolvable pairs of signals will be such that the mapping $(f, g) \mapsto f * g$ is continuous.

  As we shall see, the continuity of (f, g) is determined by giving a bound like $\|f * g\| \le C\|f\|\|g\|$, where we are intentionally vague about what (semi)norms we are using.

Convolution

- We consider four cases of continuity of convolution, corresponding to four cases of convolvable pairs of signals given above.
- For signals with support in $\mathbb{R}$, we use the usual L^p-norms.
- For signals with support in $\mathbb{R}_{\ge 0}$ we shall make use of the following fact:

  A sequence $(f_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{R}_{\ge 0};\mathbb{F})$ converges to zero if and only if, for every $T \in \mathbb{R}_{>0}$, the sequence $(f_j|[0, T])_{j\in\mathbb{Z}_{>0}}$ converges to zero in $\mathsf{L}^p([0, T];\mathbb{F})$.
- This amounts to saying that we can replace arbitrary compact subintervals $\mathbb{K} \subseteq \mathbb{R}_{\ge 0}$ with the particular compact subintervals $[0, T]$, $T \in \mathbb{R}_{>0}$.
- We have the associated seminorms $\|\cdot\|_{[0,T],p}$.

Convolution

- We then have that the mapping $(f, g) \mapsto f * g$ is continuous in the following four cases:
  1. $f, g \in \mathsf{L}^1(\mathbb{R}; \mathbb{F})$: continuity is determined by the inequality
  $$\|f * g\|_1 \leq \|f\|_1 \|g\|_1;$$
  2. $f \in \mathsf{L}^p(\mathbb{R}; \mathbb{F})$ and $g \in \mathsf{L}^q(\mathbb{R}; \mathbb{F})$ with $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$: continuity is determined by ***Young's inequality***
  $$\|f * g\|_r \leq \|f\|_p \|g\|_q;$$
  3. $f, g \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0}; \mathbb{F})$: continuity is determined by the inequality
  $$\|f * g\|_{[0,T],1} \leq \|f\|_{[0,T],1} \|g\|_{[0,T],1};$$
  4. $f \in \mathsf{L}^p(\mathbb{R}_{\geq 0}; \mathbb{F})$ and $g \in \mathsf{L}^q(\mathbb{R}_{\geq 0}; \mathbb{F})$ with $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$: continuity is determined by the inequality
  $$\|f * g\|_{[0,T],r} \leq \|f\|_{[0,T],p} \|g\|_{[0,T],q}.$$

Convolution

- Next we consider discrete-time convolution
$$f * g(k\Delta) = \sum_{j=-\infty}^{\infty} f(k\Delta - j\Delta) g(j\Delta).$$

Definition

A pair $(f, g), f, g \in \mathbb{F}^{\mathbb{Z}(\Delta)}$, is ***convolvable*** if, for each $k \in \mathbb{Z}$,
$$j\Delta \mapsto f(k\Delta - j\Delta) g(j\Delta)$$
is in $\ell^1(\mathbb{Z}(\Delta); \mathbb{F})$.

Convolution

- Discrete-time convolution has the following general properties.
  1. If (f,g) is convolvable, then (g,f) is convolvable and $f*g = g*f$.
  2. If (f,g) and (f,h) are convolvable, then $(f,g+h)$ is convolvable and $f*(f+h) = f*g+f*h$.
  3. It is generally not true that
  $$(f*g)*h = f*(g*h).$$
  4. There exists $u \in \mathbb{F}^{\mathbb{Z}(\Delta)}$ such that
  $$u*f = f, \qquad f \in \mathbb{F}^{\mathbb{Z}(\Delta)}.$$
  Thus u is a multiplicative unit for the convolution product. The signal u is easily defined explicitly, and we leave this as an exercise.
  5. For $f \in \mathbb{F}^{\mathbb{Z}(\Delta)}$, denote by
  $$\sigma(f) = \inf \operatorname{supp}(f).$$
  If $f,g \in \mathbb{F}^{\mathbb{Z}(\Delta)}$ satisfy $\sigma(f), \sigma(g) > -\infty$, then (f,g) is convolvable and
  $$f*g(k\Delta) = \begin{cases} \sum_{j=\sigma(g)/\Delta}^{k-\sigma(f)/\Delta} f(k\Delta - j\Delta)g(j\Delta), & k\Delta \geq \sigma(f)+\sigma(g), \\ 0, & k\Delta < \sigma(f)+\sigma(g). \end{cases}$$

Convolution

  6. Specialising to the case of $\sigma(f), \sigma(g) \geq 0$,
  $$f*g(k\Delta) = \begin{cases} \sum_{j=0}^{k} f(k\Delta - j\Delta)g(j\Delta), & k \geq 0, \\ 0, & k < 0. \end{cases}$$

- We now give some instances of convolvable pairs of discrete-time signals, essentially mirroring the results we already have for the continuous-time case.

Theorem

If $f,g \in \ell^1(\mathbb{Z}(\Delta);\mathbb{F})$, then $f*g \in \ell^1(\mathbb{Z}(\Delta);\mathbb{F})$. Additionally:

(i) $(f*g)*h = f*(g*h), f,g,h \in \ell^1(\mathbb{Z}(\Delta);\mathbb{F})$;

(ii) the multiplicative identity u mentioned above is in $\ell^1(\mathbb{Z}(\Delta);\mathbb{F})$, which implies that the ring $\ell^1(\mathbb{Z}(\Delta);\mathbb{F})$ with the product of convolution has a unit;

(iii) there exist $f,g \in \ell^1(\mathbb{Z}(\Delta);\mathbb{F})$, both nonzero, such that $f*g = 0$, which implies that the ring $\ell^1(\mathbb{Z}(\Delta);\mathbb{F})$ is not an integral domain;

(iv) If $h \in \ell^1(\mathbb{Z}(\Delta);\mathbb{F})$, then there exists $f,g \in \ell^1(\mathbb{Z}(\Delta);\mathbb{F})$ such that $h = f*g$ (this is trivial in this case: take $f = u$ and $g = h$).

Convolution

- We now consider convolvablility of pairs of discrete-time signals in various ℓ^p-spaces.

Theorem

If $p, q, r \in [1, \infty]$ satisfies $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, if $f \in \ell^p(\mathbb{Z}(\Delta); \mathbb{F})$, and if $g \in \ell^q(\mathbb{Z}(\Delta); \mathbb{F})$, then $f * g \in \ell^r(\mathbb{Z}(\Delta); \mathbb{F})$.

Reading for Lecture 14

Material related to this lecture can be found in the following sections of the course notes:

1. Sections IV-4.1.4, IV-4.1.5, IV-4.2.7, and IV-4.2.8.

Lecture 15
Convolution (cont'd)

- Next we consider convolvability for pairs of discrete-time signals with support in
$$\mathbb{Z}_{\geq 0}(\Delta) = \{k\Delta \in \mathbb{Z}(\Delta) \mid k \geq 0\}.$$
- The situation here is simpler than the corresponding results in the continuous-time case.

Theorem

If $f, g \in \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$, then the pair (f, g) is convolvable, and $f * g \in \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$.

- Now we turn to continuity of discrete-time convolution. First we consider the case where we have no restriction on the supports of the signals.
  1. If $f, g \in \ell^1(\mathbb{Z}(\Delta); \mathbb{F})$, then
  $$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$
  2. If $f \in \ell^p(\mathbb{Z}(\Delta); \mathbb{F})$, if $g \in \ell^q(\mathbb{Z}(\Delta); \mathbb{F})$, and if $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, then
  $$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Convolution

- For signals with support in $\ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$, we simplify the description of convergence by using particular seminorms.
- For $f \in \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$ and for $N \in \mathbb{Z}_{\geq 0}$, define
$$\|f\|_{N,p} = \left(\sum_{j=0}^{N} |f(j\Delta)|^p \right)^{1/p}.$$
- Fact: a sequence $(f_j)_{j \in \mathbb{Z}_{>0}}$ in $\ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$ converges to zero if and only if $\lim_{j \to \infty} \|f_j\|_{N,p} = 0$ for every $N \in \mathbb{Z}_{\geq 0}$.
- Then we have the following.
  1. If $f, g \in \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$ then
  $$\|f * g\|_{N,1} \leq \|f\|_{N,1} \|g\|_{N,1}.$$
  2. If $f, g \in \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F})$ and if $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, then
  $$\|f * g\|_{N,r} \leq \|f\|_{N,p} \|g\|_{N,q}.$$

System theory

- We will talk about eight kinds of systems about eight kinds of systems. These are

$$\{\text{systems, linear systems}\} \times \{\text{state-space, input/output}\} \times \{\text{continuous-time, discrete-time}\}.$$

- We shall start with the more general classes of systems, then move to the more structured classes of systems, about which we will be able to say more.

Continuous-time state space systems

Definition

A ***continuous-time state space system*** is a sextuple $(X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$, where

(i) $X \subseteq \mathbb{R}^n$ is open (the ***state space***),

(ii) $U \subseteq \mathbb{R}^m$ (the ***input-value space***),

(iii) $\mathbb{T} \subseteq \mathbb{R}$ is a continuous time-domain (the ***time-domain***),

(iv) $\mathscr{U} \subseteq U^{(\mathbb{T})}$ (the set of ***inputs***),

(v) $\boldsymbol{f} : \mathbb{T} \times X \times U \to \mathbb{R}^n$ (the ***dynamics***), and

(vi) $\boldsymbol{h} : \mathbb{T} \times X \times U \to \mathbb{R}^k$ (the ***output map***).

- In order to see the manner in which this is a system, we need a couple of definitions.

Continuous-time state space systems

Definition

A ***controlled trajectory*** of a continuous-time state space system $\sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ is a pair $(\boldsymbol{\xi}, \boldsymbol{\mu})$, where $\boldsymbol{\mu} \in \mathscr{U}$ is defined on $\mathbb{T}' \subseteq \mathbb{T}$ and where $\boldsymbol{\xi}\colon \mathbb{T}' \to X$ is locally absolutely continuous and satisfies

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).$$

Definition

A ***controlled output*** for σ is a pair $(\boldsymbol{\eta}, \boldsymbol{\mu})$, where $\boldsymbol{\mu} \in \mathscr{U}$ is defined on $\mathbb{T}' \subseteq \mathbb{T}$ and $\boldsymbol{\eta}$ satisfies

$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$

and where $(\boldsymbol{\xi}, \boldsymbol{\mu})$ is a controlled trajectory.

- Ctraj(Σ) denotes the set of controlled trajectories.
- Cout(Σ) denotes the set of controlled outputs.

Continuous-time state space systems

- Thus a controlled output $(\boldsymbol{\eta}, \boldsymbol{\mu})$ satisfies

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$
$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).$$

- Thus one obtains controlled outputs by a two-step process:
  1. given the input, determine a controlled trajectory by solving the differential equation
  $$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t));$$
  2. determine the output from
  $$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).$$
- A special case we shall study in detail later is given by

$$\dot{\xi}(t) = \mathsf{A} \circ \xi(t) + \mathsf{B} \circ \mu(t),$$
$$\eta(t) = \mathsf{C} \circ \xi(t) + \mathsf{D} \circ \mu(t),$$

for linear maps A, B, C, and D.

Continuous-time state space systems

Questions

1. What properties should the set $\mathscr{U}$ of inputs have?
2. Given properties of $\mathscr{U}$, what properties should f have so that controlled trajectories exist?
3. Continuous-time state space systems are examples of general time systems. As general time systems, what properties do they have?
4. We will care about continuity of the map $\mu \mapsto \eta$.

Reading for Lecture 15

Material related to this lecture can be found in the following sections of the course notes:

1. Sections IV-4.2.7, IV-4.2.8, and IV-4.2.9.
2. Section V-6.1.1.

Lecture 16
Continuous-time state space systems (cont'd)

- Let us give a few particular properties for continuous-time state space systems.

Definition

Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a continuous-time state space system. It is

(i) ***dynamically autonomous*** if $\boldsymbol{f}$ is independent of t, i.e., there exists $\boldsymbol{f}_0 : X \times U \to \mathbb{R}^n$ such that $\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{f}_0(\boldsymbol{x}, \boldsymbol{u})$,

(ii) ***output autonomous*** if $\boldsymbol{h}$ is independent of t, i.e., there exists $\boldsymbol{h}_0 : X \times U \to \mathbb{R}^n$ such that $\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{h}_0(\boldsymbol{x}, \boldsymbol{u})$,

(iii) ***autonomous*** if both dynamically and output autonomous, and

(iv) ***proper*** if $\boldsymbol{h}$ is independent of input, i.e., there exists $\boldsymbol{h}_0 : \mathbb{T} \times X \to \mathbb{R}^k$ such that $\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{h}_0(t, \boldsymbol{x})$.

Continuous-time state space systems

- The "autonomous" terminology mirrors that for ordinary differential equations.
- The terminology "proper" is borrowed from the theory of stationary linear systems, where it has to do with properties of the transfer function.
- Now let us consider the matter of inputs for continuous-time state space systems.
- *Notation:*
  1. $\mathsf{L}^p_{\mathsf{loc}}((\mathbb{T}); U) = \{\boldsymbol{\mu} \in U^{(\mathbb{T})} \mid \boldsymbol{\mu} \in \mathsf{L}^p_{\mathsf{loc}}(\mathrm{dom}(\boldsymbol{\mu}); U)\}$;
  2. $\mathsf{C}^0((\mathbb{T}); U) = \{\boldsymbol{\mu} \in U^{(\mathbb{T})} \mid \boldsymbol{\mu} \in \mathsf{C}^0(\mathrm{dom}(\boldsymbol{\mu}); U)\}$.
- We do this in such a manner as to ensure that, for an input $\boldsymbol{\mu} \in \mathscr{U}$, the mapping

$$(t, \boldsymbol{x}) \mapsto \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\mu}(t))$$

  satisfies the hypotheses for existence and uniqueness of solutions for ordinary differential equations.
- The reason for this, of course, is that controlled trajectories $(\boldsymbol{\xi}, \boldsymbol{\mu})$ satisfies

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).$$

Continuous-time state space systems

Example

Let us see why requiring that inputs be in $\mathsf{L}^1_{\mathsf{loc}}$ is not enough. Consider the continuous-time state space system with $X = \mathbb{R}$, $U = \mathbb{R}$, $\mathbb{T} = \mathbb{R}$, and $f(t,x,u) = u^2$. If

$$\mu(t) = \begin{cases} t^{-1/2}, & t \in \mathbb{R}_{>0}, \\ 0, & t \in \mathbb{R}_{\leq 0}. \end{cases}$$

We have $\mu \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{R};\mathbb{R})$ but

$$f(t,x,\mu(t)) = \begin{cases} t^{-1}, & t \in \mathbb{R}_{>0}, \\ 0, & t \in \mathbb{R}_{\leq 0} \end{cases}$$

is not locally integrable. Thus being in $\mathsf{L}^1_{\mathsf{loc}}$ is not, in general, an adequate property for inputs. We see, however, that if $\mu \in \mathsf{L}^\infty_{\mathsf{loc}}(\mathbb{R};\mathbb{R})$, then, for this example, things work out. Indeed, this will be enough for the general result we now state.

Continuous-time state space systems

Theorem

Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a continuous-time state space system and suppose that

(i) $\mathscr{U} \subseteq \mathsf{L}^\infty_{\mathsf{loc}}((\mathbb{T}); U)$,

(ii) for fixed $(\boldsymbol{x},\boldsymbol{u}) \in X \times U$, $t \mapsto \boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{u})$ is locally integrable,

(iii) for fixed $(t,\boldsymbol{u}) \in \mathbb{T} \times U$, $\boldsymbol{x} \mapsto \boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{u})$ is locally Lipschitz,

(iv) for fixed $t \in \mathbb{T}$, $(\boldsymbol{x},\boldsymbol{u}) \mapsto \boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{u})$ is continuous, and

(v) for $(t_0,\boldsymbol{x}_0,\boldsymbol{u}_0) \in \mathbb{T} \times X \times U$, there exists $\rho, r_1, r_2 \in \mathbb{R}_{>0}$ and $g_1, g_2 \in \mathsf{L}^1([t_0-\rho, t_0+\rho]; \mathbb{R}_{\geq 0})$ such that

(a) $\|\boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{u})\| \leq g_1(t)$ for all $t \in [t_0-\rho,t_0+\rho]$, for all $\boldsymbol{x} \in \mathsf{B}(r_1,\boldsymbol{x}_0)$ and for all $\boldsymbol{u} \in \mathsf{B}(r_2,\boldsymbol{u}_0) \cap U$, and

(b) $\|\boldsymbol{f}(t,\boldsymbol{x}_1,\boldsymbol{u}) - \boldsymbol{f}(t,\boldsymbol{x}_2,\boldsymbol{u})\| \leq g_2(t)\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$ for all $t \in [t_0-\rho,t_0+\rho]$, for all $\boldsymbol{x}_1,\boldsymbol{x}_2 \in \mathsf{B}(r_1,\boldsymbol{x}_0)$ and for all $\boldsymbol{u} \in \mathsf{B}(r_2,\boldsymbol{u}_0) \cap U$.

Then, for $\boldsymbol{\mu} \in \mathscr{U}$, the mapping $(t,\boldsymbol{x}) \mapsto \boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{\mu}(t))$ satisfies the conditions for existence and uniqueness for solutions of ordinary differential equations.

Continuous-time state space systems

- If the conditions for the theorem are satisfied for all $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$ and $\boldsymbol{\mu} \in \mathscr{U}$, there is a unique solution to the initial value problem

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)), \quad \boldsymbol{\xi}(t_0) = \boldsymbol{x}_0 \tag{3}$$

  defined for t near t_0.
- If $t_0 \in \mathbb{T}$, $\boldsymbol{x}_0 \in X$, $\mu \in \mathscr{U}$, with $t_0 \in \operatorname{dom}(\boldsymbol{\mu})$, we have a largest subinterval $I_\Sigma(t_0, \boldsymbol{x}_0, \boldsymbol{\mu})$ on which solutions of the initial value problem (3) are defined.
- The ***domain*** of Σ is

$$D_\Sigma = \{(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}) \mid t \in I_\Sigma(t_0, \boldsymbol{x}_0, \boldsymbol{\mu})\}.$$

- The ***flow*** for Σ is $\Phi^\Sigma \colon D_\Sigma \to X$ is defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^\Sigma(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}) = \boldsymbol{f}(t, \Phi^\Sigma(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t)), \quad \Phi^\Sigma(t_0, t_0, \boldsymbol{x}_0) = \boldsymbol{x}_0.$$

- Similarly with the flow and solutions for ordinary differential equations, the flow encodes all controlled trajectories for Σ.

Continuous-time state space systems

- Note that there are no conditions required for $\boldsymbol{h}$ to give the output

$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \Phi^\Sigma(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t)).$$

  But, when we consider the continuity of the input/output map we will need conditions on $\boldsymbol{h}$.
- Continuous-time state space systems are examples of general time systems. Let us consider the general time system properties of continuous-time state space systems.
  1. *They are causal:* Let $t_0 \in \mathbb{T}$. Let $(\boldsymbol{\eta}, \boldsymbol{\mu})$ be a controlled output with $t_0 \in \operatorname{dom}(\boldsymbol{\mu})$. Then the associated controlled trajectory $(\boldsymbol{\xi}, \boldsymbol{\mu})$ satisfies

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \implies \boldsymbol{\xi}(t) = \boldsymbol{\xi}(t_0) + \int_{t_0}^{t} \boldsymbol{f}(\tau, \boldsymbol{\xi}(\tau), \boldsymbol{\mu}(\tau))\,\mathrm{d}\tau.$$

     Thus $\boldsymbol{\xi}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t]$. Since

$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$

     we conclude that $\boldsymbol{\eta}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t]$.

Reading for Lecture 16

1. Sections V-6.1.1 and V-6.1.2.

Lecture 17
Continuous-time state space systems (cont'd)

2. *They are sometimes strongly causal:* If Σ is proper, then we see, that in the above argument, $\boldsymbol{\xi}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t)$ and then also that $\boldsymbol{\eta}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t)$.
3. *They are sometimes stationarity:* Assuming that the input set $\mathscr{U}$ is translation-invariant, stationarity means that

$$\boldsymbol{h}(t+a, \Phi^{\Sigma}(t+a, t_0+a, \boldsymbol{x}_0, \tau^*_{t_0, t_0+a}\boldsymbol{\mu}), \tau^*_{t_0, t_0+a}\boldsymbol{\mu}(t)) = \boldsymbol{h}(t, \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t))$$

for every $a \in \mathbb{R}_{>0}$ and every $\boldsymbol{x}_0 \in X$. One can easily see that this happens if Σ is autonomous. Thus a continuous-time state space system is stationary if it is autonomous.
4. *They are strongly stationary if they are stationary:* Because one can flow backwards in time, one can show that Σ is strongly stationary if it is stationary.
5. *They are finitely observable:* Because an output is determined by (a) an input and (b) an initial state condition at time t_0, a continuous-time state space system is finitely observable from any $\tau > t_0$ and for any $t_0 \in \mathbb{T}$.

Continuous-time state space systems

- If $U \subseteq \mathbb{R}^m$ and if $\boldsymbol{u} \in U$, then we can write $\boldsymbol{u} = (u_1, \dots, u_m)$.

Definition

A continuous-time state space system $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ is ***control-affine*** if there exists $\boldsymbol{f}_0, \boldsymbol{f}_1, \dots, \boldsymbol{f}_m \colon \mathbb{T} \times X \to$ and $\boldsymbol{h}_0, \boldsymbol{h}_1, \dots, \boldsymbol{h}_m \colon \mathbb{T} \times X \to$ such that

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{f}_0(t, \boldsymbol{x}) + \sum_{a=1}^{m} u_a \boldsymbol{f}_a(t, \boldsymbol{x})$$

and

$$\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{h}_0(t, \boldsymbol{x}) + \sum_{a=1}^{m} u_a \boldsymbol{h}_a(t, \boldsymbol{x}).$$

- (An affine function has the form "linear + constant.")
- We call $\boldsymbol{f}_0$ the ***drift dynamics*** and $\boldsymbol{f}_1, \dots, \boldsymbol{f}_m$ the ***control dynamics***.
- We call $\boldsymbol{h}_0$ the ***drift/output map*** and $\boldsymbol{h}_1, \dots, \boldsymbol{h}_m$ the ***control/output map***.

Continuous-time state space systems

- If $(\boldsymbol{\xi}, \boldsymbol{\mu}) \in \mathrm{Ctraj}(\Sigma)$ with $(\boldsymbol{\eta}, \boldsymbol{\mu})$ the corresponding controlled output, then

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}_0(t, \boldsymbol{\xi}(t)) + \sum_{a=1}^{m} \mu_a(t) \boldsymbol{f}_a(t, \boldsymbol{\xi}(t)),$$
$$\boldsymbol{\eta}(t) = \boldsymbol{h}_0(t, \boldsymbol{\xi}(t)) + \sum_{a=1}^{m} \mu_a(t) \boldsymbol{h}_a(t, \boldsymbol{\xi}(t)).$$

- We see that, if the input $\boldsymbol{\mu}$ is zero, then $\boldsymbol{\xi}$ satisfies the ordinary differential equation

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}_0(t, \boldsymbol{\xi}(t));$$

thus, in the absence of input, trajectories are solutions of the drift dynamics.

Continuous-time state space systems

- If a control-affine continuous-time state space system is dynamically autonomous, then we can enlarge our class of inputs from $\mathscr{U} \subseteq \mathrm{L}^{\infty}_{\mathrm{loc}}((\mathbb{T}); U)$ to $\mathscr{U} \subseteq \mathrm{L}^{1}_{\mathrm{loc}}((\mathbb{T}); U)$ since, in this case, we are assured that, for fixed $\boldsymbol{x}$, the function

$$t \mapsto \boldsymbol{f}_0(\boldsymbol{x}) + \sum_{a=1}^{m} \mu_a(t) \boldsymbol{f}_a(\boldsymbol{x})$$

  is locally integrable.
- Control-affine continuous-time state space systems have particular conditions in order for $\boldsymbol{f}$ to satisfy the conditions of the theorem of Slide 110.
  1. In the general case, we ask that $(t, \boldsymbol{x}) \mapsto \boldsymbol{f}_a(t, \boldsymbol{x})$, $a \in \{0, 1, \dots, m\}$, satisfy the conditions from the theorem in Slide 29 for existence and uniqueness of solutions for ordinary differential equations.
  2. In the dynamically autonomous case, we ask that $\boldsymbol{x} \mapsto \boldsymbol{f}_a(\boldsymbol{x})$, $a \in \{0, 1, \dots, m\}$, be locally Lipschitz.

Continuous-time state space systems

- One almost always works with autonomous control-affine continuous-time state space systems.
- A particularly interesting class of such systems that we will look at in detail later is the class of linear systems:

$$\begin{aligned}\dot{\xi}(t) &= \mathsf{A} \circ \xi(t) + \mathsf{B} \circ \mu(t), \\ \eta(t) &= \mathsf{C} \circ \xi(t) + \mathsf{D} \circ \mu(t),\end{aligned}$$

  for linear maps A, B, C, and D.

Reading for Lecture 17

1. Sections V-6.1.1 and V-6.1.3.

Lecture 18
Continuous-time input/output systems

- Unlike continuous-time state space systems which produce outputs by first determining a controlled trajectory, a continuous-time input/output system directly produces an output from an input.
- We will work with classes of inputs and outputs selected from the following collections of partially defined S-valued functions (S is a subset of some Euclidean space):
  1. $\mathsf{C}^0((\mathbb{T}); S) = \{f \colon \mathbb{S} \to S \mid \mathbb{S} \subseteq \mathbb{T}, f \in \mathsf{C}^0(\mathbb{S}; S)\}$;
  2. $\mathsf{L}^p_{\mathrm{loc}}((\mathbb{T}); S) = \{f \colon \mathbb{S} \to S \mid \mathbb{S} \subseteq \mathbb{T}, f \in \mathsf{L}^p_{\mathrm{loc}}(\mathbb{S}; S)\}$, $p \in [1, \infty]$.
- If $\mathscr{S}$ is a subset of one of these collections of partially defined signals and if $\mathbb{S} \subseteq \mathbb{T}$, then we denote

$$\mathscr{S}(\mathbb{S}) = \{f \in \mathscr{S} \mid \operatorname{dom}(f) = \mathbb{S}\}.$$

Continuous-time input/output systems

- We shall require a notion of convergence in a space $\mathscr{S}$ of partially defined signals.

Definition

Let $S \subseteq \mathbb{R}^n$, let $\mathbb{T}$ be a continuous time-domain, and let $\mathscr{S}$ be a subset of either $\mathsf{C}^0((\mathbb{T}); S)$ or $\mathsf{L}^p_{\mathrm{loc}}((\mathbb{T}); S)$, $p \in [1, \infty]$. A sequence $(\boldsymbol{f}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{S}$ ***converges*** to $\boldsymbol{f} \in \mathscr{S}$ if

(i) there exists a subinterval $\mathbb{S} \subseteq \mathbb{T}$ such that

$$\mathrm{dom}(\boldsymbol{f}) = \mathbb{S},\ \mathrm{dom}(\boldsymbol{f}_j) = \mathbb{S}, \quad j \in \mathbb{Z}_{>0},$$

and

(ii) $(\boldsymbol{f}_j)_{j\in\mathbb{Z}_{>0}}$ converges to $\boldsymbol{f}$ in $\mathsf{C}^0(\mathbb{S}; S)$ or $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{S}; S)$, $p \in [1, \infty]$, respectively.

Continuous-time input/output systems

Definition

A ***continuous-time input/output system*** is a 5-tuple $\Sigma = (U, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \boldsymbol{g})$ with

(i) $U \subseteq \mathbb{R}^m$,

(ii) $\mathbb{T}$ a continuous time-domain,

(iii) $\mathscr{U}$ is a subset of either $\mathsf{C}^0((\mathbb{T}); U)$ or $\mathsf{L}^p_{\mathrm{loc}}((\mathbb{T}); U)$,

(iv) $\mathscr{Y}$ is a subset of either $\mathsf{C}^0((\mathbb{T}); \mathbb{R}^k)$ or $\mathsf{L}^p_{\mathrm{loc}}((\mathbb{T}); \mathbb{R}^k)$, and

(v) $\boldsymbol{g}\colon \mathscr{U} \to \mathscr{Y}$ satisfies

(a) for a subinterval $\mathbb{S} \subseteq \mathbb{T}$, the restriction $\boldsymbol{g}_{\mathbb{S}} = \boldsymbol{g}|\mathscr{U}(\mathbb{S})$ takes values in $\mathscr{Y}(\mathbb{S})$,

(b) for subintervals $\mathbb{S}' \subseteq \mathbb{S}$,

$$\boldsymbol{g}_{\mathbb{S}}(\boldsymbol{\mu})|\mathbb{S}' = \boldsymbol{g}_{\mathbb{S}'}(\boldsymbol{\mu}|\mathbb{S}'), \quad \boldsymbol{\mu} \in \mathscr{U}(\mathbb{S}),$$

and

(c) for any subinterval $\mathbb{S} \subseteq \mathbb{T}$, $\boldsymbol{g}_{\mathbb{S}}$ is continuous (in the sense that it maps convergent sequences to convergent sequences).

Continuous-time input/output systems

- While we will encounter an important class of continuous-time input/output systems where we will require the possibility of partially defined spaces of inputs and outputs, it is common for inputs and outputs to be defined on all of $\mathbb{T}$, i.e., $\mathscr{U}(\mathbb{S}) = \varnothing$ and $\mathscr{Y}(\mathbb{S}) = \varnothing$ unless $\mathbb{S} = \mathbb{T}$.
- In such cases, the properties (v)(a) and (v)(b) having to do with restriction are moot, and so the definition simplifies significantly.
- Most interesting classes of input/output systems one encounters in practice are continuous. And if they are not, they should be.

Continuous-time input/output systems

- Why do we care about continuity?
  1. Some problems in system theory are solved by first finding a sequence of approximating solutions, and then taking a limit. Without continuity, such constructions are meaningless.
  2. In "real life" applications of system theory, the models with which one works are seldom perfectly accurate, the implementation of solutions of system-theoretic problems is seldom faithful. In such cases, continuity will ensure that the idealised solution will still work reasonably well on the actual system.
  3. Optimisation is a common problem in system theory. Thus one wishes to accomplish a system theoretic task while minimising a cost function. Continuity is frequently an important ingredient for ensuring the existence of optimal solutions, e.g., continuous functions on compact sets achieve a minimum.
  4. Certain kinds of stability in system theory can be phrased as continuity with respect to certain topologies on the spaces of inputs and outputs. For example, so-called bounded-input/bounded-output stability for linear systems means continuity of the system when inputs and outputs are equipped with the L^∞-norm.

Continuous-time input/output systems

- Continuous-time input/output systems are examples of general time systems. As such they are susceptible to possessing attributes of general time systems.
- But these systems have their own definitions of causality that are different from those for continuous-time state space systems as a result of there being no natural initial time t_0.

Definition

A continuous-time input/output system is:

(i) ***causal*** if, for every $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathscr{U}$ with $\mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$ and for every $t \in \mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$,

$$\boldsymbol{\mu}_1|(\mathbb{T}_{\leq t} \cap \mathrm{dom}(\boldsymbol{\mu}_1)) = \boldsymbol{\mu}_2|(\mathbb{T}_{\leq t} \cap \mathrm{dom}(\boldsymbol{\mu}_2)) \quad \Longrightarrow \quad \boldsymbol{g}(\boldsymbol{\mu}_1)(t) = \boldsymbol{g}(\boldsymbol{\mu}_2)(t);$$

(ii) ***strongly causal*** if, for every $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathscr{U}$ with $\mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$ and for every $t \in \mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$,

$$\boldsymbol{\mu}_1|(\mathbb{T}_{< t} \cap \mathrm{dom}(\boldsymbol{\mu}_1)) = \boldsymbol{\mu}_2|(\mathbb{T}_{< t} \cap \mathrm{dom}(\boldsymbol{\mu}_2)) \quad \Longrightarrow \quad \boldsymbol{g}(\boldsymbol{\mu}_1)(t) = \boldsymbol{g}(\boldsymbol{\mu}_2)(t).$$

Continuous-time input/output systems

Definition

A continuous-time input/output system with $\sup \mathbb{T} = \infty$ is:

(i) ***stationary*** if $\tau_a^*(\mathscr{U}) \subseteq \mathscr{U}$ for every $a \in \mathbb{R}_{>0}$ and if, for every $\boldsymbol{\mu} \in \mathscr{U}$,

$$\boldsymbol{g}(\tau_a^* \boldsymbol{\mu}) = \tau_a^* \boldsymbol{g}(\boldsymbol{\mu});$$

(ii) ***strongly stationary*** if it is stationary and if, for every $a \in \mathbb{R}_{>0}$ and every $\boldsymbol{\mu} \in \mathscr{U}$, there exists $\boldsymbol{\mu}' \in \mathscr{U}$ such that

$$\boldsymbol{g}(\boldsymbol{\mu}) = \boldsymbol{g}(\tau_a^* \boldsymbol{\mu}').$$

- While the definitions differ in detail from those used for continuous-time state space systems, the essence is the same:
  1. causality means that outputs do not depend on future values of inputs;
  2. stationarity means invariance under time shift.

Continuous-time input/output systems

- Given these definitions, we have the following properties of continuous-time input/output systems.
  1. *They are generally not causal:* As an example, take $U = \mathbb{R}$, $\mathbb{T} = \mathbb{R}$, $\mathscr{U} = \mathscr{Y} = \mathrm{C}^0(\mathbb{R};\mathbb{R})$ (no partially defined inputs or outputs), and $g(\mu)(t) = \mu(-t)$ or $g(\mu)(t) = \mu(t+a)$ for $a \in \mathbb{R}_{>0}$. Indeed, it is pretty easy to build noncausal systems.
  2. *They are generally not stationary:* As an example, take $U = \mathbb{R}$, $\mathbb{T} = \mathbb{R}$, $\mathscr{U} = \mathrm{L}^1(\mathbb{R};\mathbb{R})$, $\mathscr{Y} = \mathrm{C}^0(\mathbb{R};\mathbb{R})$, and

$$g(\mu)(t) = \int_{-\infty}^{t} \sin(\tau)\mu(\tau)\,\mathrm{d}\tau.$$

     It is easy to directly verify that, for arbitrary $a \in \mathbb{R}_{>0}$, $\tau_a^* g(\mu) \neq g(\tau_a^* \mu)$.
  3. *They are finitely observable:* This is just because they are functional as general time systems.

Reading for Lecture 18

1. Sections V-6.2.1 and V-6.2.2.

Lecture 19
Continuous-time input/output systems from continuous-time state space systems

- Let $\Sigma = (X, U, Y, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \boldsymbol{f}, \boldsymbol{h})$ be a continuous-time state space system and let $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$.
- Then we can define a continuous-time input/output system $\Sigma_{\mathrm{i/o}}(t_0, \boldsymbol{x}_0) = (U, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \boldsymbol{g})$ by

$$\boldsymbol{g}(\boldsymbol{\mu})(t) = \boldsymbol{h}(t, \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t)).$$

  where, we recall that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}) = \boldsymbol{f}(t, \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t)), \quad \Phi^{\Sigma}(t_0, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}) = \boldsymbol{x}_0.$$

- We assume that $\boldsymbol{f}$ satisfies the previous conditions giving existence and uniqueness of trajectories for the existence and uniqueness of controlled trajectories.

Continuous-time input/output systems from continuous-time state space systems

- We can then easily verify that $\boldsymbol{g}$ has all the properties of a continuous-time input/output systems, except possibly continuity.
- To verify continuity, we do two things:
  1. give appropriate properties of $\boldsymbol{h}$;
  2. select an appropriate set of outputs based on the properties of $\boldsymbol{h}$.
- We shall state a theorem that considers various special cases. The fully rigorous proof is difficult. Instead we shall give some reasons for why the conditions are as they are by examining the system equations

$$\begin{aligned}\dot{\boldsymbol{\xi}}(t) &= \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),\\ \boldsymbol{\eta}(t) &= \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).\end{aligned} \tag{4}$$

  We shall use the fact that, in general, inputs are in $\mathsf{L}^{\infty}_{\mathrm{loc}}((\mathbb{T}); U)$ and that trajectories ξ are locally absolutely continuous, and so continuous, as functions of time.

Continuous-time input/output systems from continuous-time state space systems

Theorem

Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a continuous-time state space system, let $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$, and let $\Sigma_{\mathrm{i/o}}$ be as above. Then, under the following sets of conditions, $\Sigma_{\mathrm{i/o}}$ is a continuous-time input/output system.

(i) ***The most general case:***

(a) $U \subseteq \mathbb{R}^m$ is locally compact,

(b) the map $t \mapsto \boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u})$ is measurable for each $(\boldsymbol{x}, \boldsymbol{u}) \in X \times U$,

(c) the map $(\boldsymbol{x}, \boldsymbol{u}) \mapsto \boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u})$ is continuous for each $t \in \mathbb{T}$, and

(d) for each $(t, \boldsymbol{x}, \boldsymbol{u}) \in \mathbb{T} \times X \times U$, there exist $r_1, r_2, \alpha \in \mathbb{R}_{>0}$ and

$$g \in \mathsf{L}^1([t-\alpha, t+\alpha]; \mathbb{R}_{\geq 0})$$

such that

$$\|\boldsymbol{h}(s, \boldsymbol{x}', \boldsymbol{u}')\| \leq g(s), \quad (s, \boldsymbol{x}', \boldsymbol{u}') \in ([t-\alpha, t+\alpha] \cap \mathbb{T}) \times \mathsf{B}^n(r_1, \boldsymbol{x}) \times \mathsf{B}^n(r_2, \boldsymbol{u}),$$

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(e) $\mathscr{U} = \mathsf{L}^\infty_{\mathrm{loc}}((\mathbb{T}); U)$, and

(f) $\mathscr{Y} = \mathsf{L}^1_{\mathrm{loc}}((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Referring to (4), the conditions on $\boldsymbol{h}$ ensure that

1. the t-dependence of $\boldsymbol{h}$ is locally integrable;
2. the t-dependence of $\boldsymbol{h}$ on $\boldsymbol{\xi}(t)$ is locally bounded ($t \mapsto \boldsymbol{\xi}(t)$ is continuous and $\boldsymbol{h}$ is continuous in $\boldsymbol{x}$);
3. the t-dependence of $\boldsymbol{h}$ on $\boldsymbol{\mu}(t)$ is locally bounded ($t \mapsto \boldsymbol{\mu}(t)$ is locally bounded and $\boldsymbol{h}$ is continuous in $\boldsymbol{u}$).

Since continuous combinations of locally integrable functions and locally bounded functions will be locally integrable, we deduce that $\mathscr{Y} = \mathsf{L}^1_{\mathrm{loc}}((\mathbb{T}); \mathbb{R}^k)$. □

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(ii) ***The output autonomous case:***

(a) $U \subseteq \mathbb{R}^m$ *is locally compact,*

(b) Σ *is output autonomous,*

(c) *the map* $(\boldsymbol{x}, \boldsymbol{u}) \mapsto \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{u})$ *is continuous,*

(d) $\mathscr{U} \subseteq \mathsf{L}^{\infty}_{\mathsf{loc}}((\mathbb{T}); U)$*, and*

(e) $\mathscr{Y} = \mathsf{L}^{\infty}_{\mathsf{loc}}((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Similar to Case (i), except that $\boldsymbol{h}$ *is independent of* t*, and so we only have local boundedness of output.* □

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(iii) ***The output autonomous, proper case:***

(a) Σ *is output autonomous and proper,*

(b) *the map* $\boldsymbol{x} \mapsto \boldsymbol{h}(\boldsymbol{x})$ *is continuous,*

(c) $\mathscr{U} \subseteq \mathsf{L}^{\infty}_{\mathsf{loc}}((\mathbb{T}); U)$*, and*

(d) $\mathscr{Y} = \mathsf{C}^0((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Referring to (4), we see that $t \mapsto \boldsymbol{\xi}(t)$ *is continuous, and so the output is continuous since* $\boldsymbol{h}$ *depends continuously on* $\boldsymbol{x}$. □

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(iv) ***The general control-affine case:***

(a) Σ *is control-affine,*

(b) *the maps* $t \mapsto \boldsymbol{h}_a(t, \boldsymbol{x})$, $a \in \{0, 1, \dots, m\}$, *are measurable for each* $\boldsymbol{x} \in X$,

(c) *the maps* $\boldsymbol{x} \mapsto \boldsymbol{h}_a(t, \boldsymbol{x})$, $a \in \{0, 1, \dots, m\}$, *are continuous for each* $t \in \mathbb{T}$, *and*

(d) *for each* $(t, \boldsymbol{x}) \in \mathbb{T} \times X$, *there exist* $r, \alpha \in \mathbb{R}_{>0}$ *and*

$$g \in \mathsf{L}^1_{\mathrm{loc}}([t - \alpha, t + \alpha]; \mathbb{R}_{\geq 0})$$

such that

$$\|\boldsymbol{h}_a(s, \boldsymbol{x}')\| \leq g(s), \quad a \in \{0, 1, \dots, m\},\ (s, \boldsymbol{x}') \in ([t-\alpha, t+\alpha] \cap \mathbb{T} \times \mathsf{B}^n(r, \boldsymbol{x}),$$

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(e) $\mathscr{U} \subseteq \mathsf{L}^\infty_{\mathrm{loc}}((\mathbb{T}); U)$, *and*

(f) $\mathscr{Y} = \mathsf{L}^1_{\mathrm{loc}}((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Just like Case (i). □

(v) ***The control-affine output autonomous case:***

(a) Σ *is control-affine and output autonomous,*

(b) *the maps* $\boldsymbol{x} \mapsto \boldsymbol{h}_a(\boldsymbol{x})$, $a \in \{0, 1, \dots, m\}$, *are continuous,*

(c) $\mathscr{U} \subseteq \mathsf{L}^\infty_{\mathrm{loc}}((\mathbb{T}); U)$, *and*

(d) $\mathscr{Y} = \mathsf{L}^\infty_{\mathrm{loc}}((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Same as Case (ii). □

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(vi) ***The control-affine output autonomous, proper case:***

(a) Σ *is control-affine, output autonomous, and proper,*

(b) *the map* $\boldsymbol{x} \mapsto \boldsymbol{h}_0(\boldsymbol{x})$ *is continuous,*

(c) $\mathscr{U} \subseteq \mathsf{L}^{\infty}_{\mathsf{loc}}((\mathbb{T}); U)$*, and*

(d) $\mathscr{Y} = \mathsf{C}^0((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Same as (iii). □

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(vii) ***The control-affine autonomous case:***

(a) Σ *is control-affine and autonomous,*

(b) *the maps* $\boldsymbol{x} \mapsto \boldsymbol{h}_a(\boldsymbol{x})$, $a \in \{0, 1, \dots, m\}$*, are continuous,*

(c) $\mathscr{U} \subseteq \mathsf{L}^{1}_{\mathsf{loc}}((\mathbb{T}); U)$*, and*

(d) $\mathscr{Y} = \mathsf{L}^{1}_{\mathsf{loc}}((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

Here we note that the differential equation part of (4) permits us to use inputs that are only locally integrable, and not subject to the stronger condition of being locally bounded. The reason for this is that $\boldsymbol{f}$ *is linear in inputs in the control-affine case. Then we see that the outputs are similarly naturally locally integrable.* □

Continuous-time input/output systems from continuous-time state space systems

Theorem (cont'd)

(viii) ***The control-affine, autonomous, proper case:***

(a) Σ *is control-affine, autonomous, and proper,*

(b) the map $\boldsymbol{x} \mapsto \boldsymbol{h}_0(\boldsymbol{x})$ *is continuous,*

(c) $\mathscr{U} \subseteq \mathsf{L}^1_{\mathsf{loc}}((\mathbb{T}); U)$*, and*

(d) $\mathscr{Y} = \mathsf{C}^0((\mathbb{T}); \mathbb{R}^k)$.

Idea of proof.

As in Case (vii), we can take inputs to be locally integrable. Then the outputs are continuous, since $t \mapsto \boldsymbol{\xi}(t)$ *is continuous and* $\boldsymbol{h}_0$ *is continuous.* □

Reading for Lecture 19

1. Section V-6.2.3.

Lecture 20
Discrete-time state space systems

- The development of discrete-time state space systems proceeds rather like that for continuous-time state space systems. So what we say will be a little repetitious. But there are differences we will point out.

Definition

A ***discrete-time state space system*** is a sextuple $(X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$, where

(i) $X \subseteq \mathbb{R}^n$ is open (the ***state space***),
(ii) $U \subseteq \mathbb{R}^m$ (the ***input-value space***),
(iii) $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ is a discrete time-domain (the ***time-domain***),
(iv) $\mathscr{U} \subseteq U^{(\mathbb{T})}$ (the set of ***inputs***),
(v) $\boldsymbol{f}\colon \mathbb{T} \times X \times U \to \mathbb{R}^n$ (the ***dynamics***), and
(vi) $\boldsymbol{h}\colon \mathbb{T} \times X \times U \to \mathbb{R}^k$ (the ***output map***).

Discrete-time state space systems

Definition

A ***controlled trajectory*** of a discrete-time state space system $\sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ is a pair $(\boldsymbol{\xi}, \boldsymbol{\mu})$, where $\boldsymbol{\mu} \in \mathscr{U}$ is defined on $\mathbb{T}' \subseteq \mathbb{T}$ and where $\boldsymbol{\xi}\colon \mathbb{T}' \to X$ satisfies

$$\boldsymbol{\xi}(t + \Delta) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).$$

Definition

A ***controlled output*** for σ is a pair $(\boldsymbol{\eta}, \boldsymbol{\mu})$, where $\boldsymbol{\mu} \in \mathscr{U}$ is defined on $\mathbb{T}' \subseteq \mathbb{T}$ and $\boldsymbol{\eta}$ satisfies

$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$

and where $(\boldsymbol{\xi}, \boldsymbol{\mu})$ is a controlled trajectory.

- Ctraj(Σ) denotes the set of controlled trajectories.
- Cout(Σ) denotes the set of controlled outputs.

Discrete-time state space systems

- Thus a controlled output $(\boldsymbol{\eta}, \boldsymbol{\mu})$ satisfies

$$\begin{aligned}\boldsymbol{\xi}(t+\Delta) &= \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),\\ \boldsymbol{\eta}(t) &= \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).\end{aligned}$$

- Thus one obtains controlled outputs by a two-step process:
  1. given the input, determine a controlled trajectory by solving the difference equation
  $$\boldsymbol{\xi}(t+\Delta) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t));$$
  2. determine the output from
  $$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)).$$
- A special case we shall study in detail later is given by

$$\begin{aligned}\xi(t+\Delta) &= \mathsf{A} \circ \xi(t) + \mathsf{B} \circ \mu(t),\\ \eta(t) &= \mathsf{C} \circ \xi(t) + \mathsf{D} \circ \mu(t),\end{aligned}$$

for linear maps A, B, C, and D.

Discrete-time state space systems

- Let us give a few particular properties for discrete-time state space systems.

Definition

Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a discrete-time state space system. It is

(i) ***dynamically autonomous*** if $\boldsymbol{f}$ is independent of t, i.e., there exists $\boldsymbol{f}_0 : X \times U \to \mathbb{R}^n$ such that $\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{f}_0(\boldsymbol{x}, \boldsymbol{u})$,

(ii) ***output autonomous*** if $\boldsymbol{h}$ is independent of t, i.e., there exists $\boldsymbol{h}_0 : X \times U \to \mathbb{R}^n$ such that $\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{h}_0(\boldsymbol{x}, \boldsymbol{u})$,

(iii) ***autonomous*** if both dynamically and output autonomous, and

(iv) ***proper*** if $\boldsymbol{h}$ is independent of input, i.e., there exists $\boldsymbol{h}_0 : \mathbb{T} \times X \to \mathbb{R}^k$ such that $\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{h}_0(t, \boldsymbol{x})$.

Discrete-time state space systems

- The “autonomous” terminology mirrors that for ordinary difference equations.
- The terminology “proper” is borrowed from the theory of stationary linear systems, where it has to do with properties of the transfer function.
- For discrete-time state space systems, one does not have to think carefully about the kind of inputs one uses. One can just take $\mathscr{U} \subseteq \ell_{\mathrm{loc}}((\mathbb{T}); U)$.
- For discrete-time systems, one does not have to think carefully about the property of $\boldsymbol{f}$ to ensure existence of trajectories. Because the trajectory is simply determined by the difference equation

$$\boldsymbol{\xi}(t+\Delta) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$

  and difference equations are simply solved by recursion, any mapping $\boldsymbol{f}$ will give the existence of trajectories.

Discrete-time state space systems

- By the same reasoning as we saw for difference equations, trajectories are not generally uniquely defined by their initial conditions. This is because of the lack of invertibility of difference equations. However, one does have forward uniqueness, meaning that the trajectory with initial state $\boldsymbol{x}_0$ at time t_0 is uniquely defined for $t \geq t_0$.
- So we can define the flow.
- The ***flow*** for Σ is $\Phi^{\Sigma} \colon D_{\Sigma} \to X$ is defined by

$$\Phi^{\Sigma}(t+\Delta, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}) = \boldsymbol{f}(t, \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t)), \quad \Phi^{\Sigma}(t_0, t_0, \boldsymbol{x}_0) = \boldsymbol{x}_0.$$

- Similarly with the flow and solutions for ordinary difference equations, the flow encodes all controlled trajectories for Σ.

Discrete-time state space systems

- Note that there are no conditions required for $\boldsymbol{h}$ to give the output
$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t)).$$
But, when we consider the continuity of the input/output map we will need conditions on $\boldsymbol{h}$.
- Discrete-time state space systems are examples of general time systems. Let us consider the general time system properties of discrete-time state space systems.
  1. *They are causal:* Let $t_0 \in \mathbb{T}$. Let $(\boldsymbol{\eta}, \boldsymbol{\mu})$ be a controlled trajectory with $t_0 \in \operatorname{dom}(\boldsymbol{\mu})$. Then a controlled trajectory $(\boldsymbol{\xi}, \boldsymbol{\mu})$ satisfies
  $$\boldsymbol{\xi}(t + \Delta) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$
  and so is solved by recursion, as we have seen. Thus $\boldsymbol{\xi}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t]$. Since
  $$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$
  we conclude that $\boldsymbol{\eta}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t]$.

Discrete-time state space systems

2. *They are sometimes strongly causal:* If Σ is proper, then we see, that in the above argument, $\boldsymbol{\xi}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t)$ and then also that $\boldsymbol{\eta}(t)$ depends only on the value of $\boldsymbol{\mu}$ on $[t_0, t)$.
3. *They are sometimes stationarity:* Assuming that the input set $\mathscr{U}$ is translation-invariant, stationarity means that
$$\boldsymbol{h}(t + a, \Phi^{\Sigma}(t + a, t_0 + a, \boldsymbol{x}_0, \tau^*_{t_0, t_0+a}\boldsymbol{\mu}), \tau^*_{t_0, t_0+a}\boldsymbol{\mu}(t)) = \boldsymbol{h}(t, \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}), \boldsymbol{\mu}(t))$$
for every $a \in \mathbb{Z}_{>0}(\Delta)$ and every $\boldsymbol{x}_0 \in X$. One can easily see that this happens if Σ is autonomous. Thus a discrete-time state space system is stationary if it is autonomous.
4. *They are not generally strongly stationary even if they are stationary:* Because one cannot flow backwards in time, it is no longer the case that Σ is strongly stationary if it is stationary. This is an important difference with the continuous-time case. (See Exercise V-6.3.2.)
5. *They are finitely observable:* Because an output is determined by (a) an input and (b) an initial state condition at time t_0, a discrete-time state space system is finitely observable from any $\tau > t_0$ and for any $t_0 \in \mathbb{T}$.

Discrete-time state space systems

- If $U \subseteq \mathbb{R}^m$ and if $\boldsymbol{u} \in U$, then we can write $\boldsymbol{u} = (u_1, \dots, u_m)$.

Definition

A discrete-time state space system $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ is ***control-affine*** if there exists $\boldsymbol{f}_0, \boldsymbol{f}_1, \dots, \boldsymbol{f}_m \colon \mathbb{T} \times X \to$ and $\boldsymbol{h}_0, \boldsymbol{h}_1, \dots, \boldsymbol{h}_m \colon \mathbb{T} \times X \to$ such that

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{f}_0(t, \boldsymbol{x}) + \sum_{a=1}^{m} u_a \boldsymbol{f}_a(t, \boldsymbol{x})$$

and

$$\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{h}_0(t, \boldsymbol{x}) + \sum_{a=1}^{m} u_a \boldsymbol{h}_a(t, \boldsymbol{x}).$$

- We call $\boldsymbol{f}_0$ the ***drift dynamics*** and $\boldsymbol{f}_1, \dots, \boldsymbol{f}_m$ the ***control dynamics***.
- We call $\boldsymbol{h}_0$ the ***drift/output map*** and $\boldsymbol{f}_1, \dots, \boldsymbol{f}_m$ the ***control/output map***.

Discrete-time state space systems

- If $(\boldsymbol{\xi}, \boldsymbol{\mu}) \in \mathrm{Ctraj}(\Sigma)$ with $(\boldsymbol{\eta}, \boldsymbol{\mu})$ the corresponding controlled output, then

$$\boldsymbol{\xi}(t + \Delta) = \boldsymbol{f}_0(t, \boldsymbol{\xi}(t)) + \sum_{a=1}^{m} \mu_a(t) \boldsymbol{f}_a(t, \boldsymbol{\xi}(t)),$$

$$\boldsymbol{\eta}(t) = \boldsymbol{h}_0(t, \boldsymbol{\xi}(t)) + \sum_{a=1}^{m} \mu_a(t) \boldsymbol{h}_a(t, \boldsymbol{\xi}(t)).$$

- We see that, if the input $\boldsymbol{\mu}$ is zero, then $\boldsymbol{\xi}$ satisfies the ordinary difference equation

$$\boldsymbol{\xi}(t + \Delta) = \boldsymbol{f}_0(t, \boldsymbol{\xi}(t));$$

thus, in the absence of input, trajectories are solutions of the drift dynamics.

Discrete-time state space systems

- One almost always works with autonomous control-affine discrete-time state space systems.
- A particularly interesting class of such systems that we will look at in detail later is the class of linear systems:

$$\begin{aligned}\xi(t+\Delta) &= \mathsf{A}\circ\xi(t) + \mathsf{B}\circ\mu(t),\\ \eta(t) &= \mathsf{C}\circ\xi(t) + \mathsf{D}\circ\mu(t),\end{aligned}$$

  for linear maps A, B, C, and D.

Reading for Lecture 20

1. Sections V-6.3.1, V-6.3.2, and V-6.3.3.

Lecture 21
Discrete-time input/output systems

- Our constructions for discrete-time input/output systems are conducted much like those for continuous-time input/output systems. So there is much repetition here.
- Unlike discrete-time state space systems which produce outputs by first determining a controlled trajectory, a discrete-time input/output system directly produces an output from an input.
- We will work with classes of inputs and outputs selected from the space $\ell_{\text{loc}}((\mathbb{T}); S)$ of partially defined S-valued functions (S is a subset of some Euclidean space).
- If $\mathscr{S}$ is a subset of one of these collections of partially defined signals and if $\mathbb{S} \subseteq \mathbb{T}$, then we denote

$$\mathscr{S}(\mathbb{S}) = \{\boldsymbol{f} \in \mathscr{S} \mid \operatorname{dom}(\mathbb{S}) = \mathbb{S}\}.$$

Discrete-time input/output systems

- We shall require a notion of convergence in a space $\mathscr{S}$ of partially defined signals.

Definition

Let $S \subseteq \mathbb{R}^n$, let $\mathbb{T}$ be a discrete time-domain, and let $\mathscr{S}$ be a subset of $\ell_{\text{loc}}((\mathbb{T}); S)$. A sequence $(\boldsymbol{f}_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{S}$ ***converges*** to $\boldsymbol{f} \in \mathscr{S}$ if

(i) there exists a subinterval $\mathbb{S} \subseteq \mathbb{T}$ such that

$$\operatorname{dom}(\boldsymbol{f}) = \mathbb{S}, \ \operatorname{dom}(\boldsymbol{f}_j) = \mathbb{S}, \quad j \in \mathbb{Z}_{>0},$$

and

(ii) $(\boldsymbol{f}_j)_{j\in\mathbb{Z}_{>0}}$ converges to $\boldsymbol{f}$ in $\ell_{\text{loc}}(\mathbb{S}; S)$.

Discrete-time input/output systems

Definition

A ***discrete-time input/output system*** is a 5-tuple $\Sigma = (U, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \boldsymbol{g})$ with

(i) $U \subseteq \mathbb{R}^m$,

(ii) $\mathbb{T}$ a discrete time-domain,

(iii) $\mathscr{U}$ is a subset of $\ell_{\text{loc}}((\mathbb{T}); U)$,

(iv) $\mathscr{Y}$ is a subset of $\ell_{\text{loc}}((\mathbb{T}); \mathbb{R}^k)$, and

(v) $\boldsymbol{g} \colon \mathscr{U} \to \mathscr{Y}$ satisfies

  (a) for a sub-time-domain $\mathbb{S} \subseteq \mathbb{T}$, the restriction $\boldsymbol{g}_{\mathbb{S}} = \boldsymbol{g}|\mathscr{U}(\mathbb{S})$ takes values in $\mathscr{Y}(\mathbb{S})$,

  (b) for sub-time-domains $\mathbb{S}' \subseteq \mathbb{S}$,

$$\boldsymbol{g}_{\mathbb{S}}(\boldsymbol{\mu})|\mathbb{S}' = \boldsymbol{g}_{\mathbb{S}'}(\boldsymbol{\mu}|\mathbb{S}'), \quad \boldsymbol{\mu} \in \mathscr{U}(\mathbb{S}),$$

  and

  (c) for any sub-time-domain $\mathbb{S} \subseteq \mathbb{T}$, $\boldsymbol{g}_{\mathbb{S}}$ is continuous (in the sense that it maps convergent sequences to convergent sequences).

Discrete-time input/output systems

- While we will encounter an important class of discrete-time input/output systems where we will require the possibility of partially defined spaces of inputs and outputs, it is common for inputs and outputs to be defined on all of $\mathbb{T}$, i.e., $\mathscr{U}(\mathbb{S}) = \varnothing$ and $\mathscr{Y}(\mathbb{S}) = \varnothing$ unless $\mathbb{S} = \mathbb{T}$.
- In such cases, the properties (v)(a) and (v)(b) having to do with restriction are moot, and so the definition simplifies significantly.
- Most interesting classes of input/output systems one encounters in practice are continuous. And if they are not, they should be.

Discrete-time input/output systems

- Discrete-time input/output systems are examples of general time systems. As such they are susceptible to possessing attributes of general time systems.
- But these systems have their own definitions of causality that are different from those for discrete-time state space systems as a result of there being no natural initial time t_0.

Definition

A discrete-time input/output system is:

(i) ***causal*** if, for every $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathscr{U}$ with $\mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$ and for every $t \in \mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$,

$$\boldsymbol{\mu}_1|(\mathbb{T}_{\leq t} \cap \mathrm{dom}(\boldsymbol{\mu}_1)) = \boldsymbol{\mu}_2|(\mathbb{T}_{\leq t} \cap \mathrm{dom}(\boldsymbol{\mu}_2)) \quad \Longrightarrow \quad \boldsymbol{g}(\boldsymbol{\mu}_1)(t) = \boldsymbol{g}(\boldsymbol{\mu}_2)(t);$$

(ii) ***strongly causal*** if, for every $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathscr{U}$ with $\mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$ and for every $t \in \mathrm{dom}(\boldsymbol{\mu}_1) = \mathrm{dom}(\boldsymbol{\mu}_2)$,

$$\boldsymbol{\mu}_1|(\mathbb{T}_{< t} \cap \mathrm{dom}(\boldsymbol{\mu}_1)) = \boldsymbol{\mu}_2|(\mathbb{T}_{< t} \cap \mathrm{dom}(\boldsymbol{\mu}_2)) \quad \Longrightarrow \quad \boldsymbol{g}(\boldsymbol{\mu}_1)(t) = \boldsymbol{g}(\boldsymbol{\mu}_2)(t).$$

Discrete-time input/output systems

Definition

A discrete-time input/output system with $\sup \mathbb{T} = \infty$ is:

(i) ***stationary*** if $\tau_a^*(\mathscr{U}) \subseteq \mathscr{U}$ for every $a \in \mathbb{Z}_{>0}(\Delta)$ and if, for every $\boldsymbol{\mu} \in \mathscr{U}$,

$$\boldsymbol{g}(\tau_a^* \boldsymbol{\mu}) = \tau_a^* \boldsymbol{g}(\boldsymbol{\mu});$$

(ii) ***strongly stationary*** if it is stationary and if, for every $a \in \mathbb{Z}_{>0}(\Delta)$ and every $\boldsymbol{\mu} \in \mathscr{U}$, there exists $\boldsymbol{\mu}' \in \mathscr{U}$ such that

$$\boldsymbol{g}(\boldsymbol{\mu}) = \boldsymbol{g}(\tau_a^* \boldsymbol{\mu}').$$

- While the definitions differ in detail from those used for discrete-time state space systems, the essence is the same:
  1. causality means that outputs do not depend on future values of inputs;
  2. stationarity means invariance under time shift.

Discrete-time input/output systems

- Given these definitions, we have the following properties of discrete-time input/output systems.
  1. *They are generally not causal:* As an example, take $U = \mathbb{R}$, $\mathbb{T} = \mathbb{Z}$, $\mathscr{U} = \mathscr{Y} = \ell_{\text{loc}}(\mathbb{Z};\mathbb{R})$ (no partially defined inputs or outputs), and $g(\mu)(t) = \mu(-t)$ or $g(\mu)(t) = \mu(t+a)$ for $a \in \mathbb{Z}_{>0}$. Indeed, it is pretty easy to build noncausal systems.
  2. *They are generally not stationary:* As an example, take $U = \mathbb{R}$, $\mathbb{T} = \mathbb{Z}$, $\mathscr{U} = \ell_{\text{loc}}(\mathbb{Z};\mathbb{R})$, and
  $$g(\mu)(t) = \sum_{-\infty}^{t} \sin(\tau)\mu(\tau)\,\mathrm{d}\tau.$$
  It is easy to directly verify that, for arbitrary $a \in \mathbb{Z}_{>0}$, $\tau_a^* g(\mu) \neq g(\tau_a^*\mu)$.
  3. *They are finitely observable:* This is just because they are functional as general time systems.

Discrete-time input/output systems

- We can then easily verify that $\boldsymbol{g}$ has all the properties of a discrete-time input/output systems, except possibly continuity.
- Unlike the continuous-time case, where we had a rather elaborate theorem giving continuity of the input/output map, in the discrete-time case things are rather simpler.

Theorem

Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a discrete-time state space system. Assume that $\boldsymbol{f}$ is continuous and that $\boldsymbol{h}$ is output autonomous and a continuous mapping from $X \times U$ to $\mathbb{R}^k$. Let $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$. Then $\Sigma_{\text{i/o}}(t_0, \boldsymbol{x}_0) = (U, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \boldsymbol{g})$, with $\boldsymbol{g}$ as defined above, defines a discrete-time input/output system.

Reading for Lecture 21

1. Sections V-6.4.1, V-6.4.2, and V-6.4.3.

Lecture 22
Linearisation

- We have introduced four classes of systems thus far.
- The remaining four classes of systems we will study are linear versions of the first four.
- Linear systems are important because
  1. they sometimes occur in nature and
  2. they arise from linearisation of systems that are not necessarily linear.
- In the next few lectures, we will carefully study linearisation.

Linearisation of continuous-time state space systems

- Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a continuous-time state space system and let $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) \in \text{Ctraj}(\Sigma)$ with $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0) \in \text{Cout}(\Sigma)$ the corresponding controlled output. Let $\text{dom}(\boldsymbol{\mu}_0) = \mathbb{T}'$.
- Call $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)$ the ***reference trajectory*** and $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0)$ the ***reference output***.
- We will linearise about this the reference. This means that we consider a controlled trajectory $(\boldsymbol{\xi}, \boldsymbol{\mu})$ that is "nearby" $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)$. We let $\boldsymbol{\nu}$, $\boldsymbol{\omega}$, and $\boldsymbol{\gamma}$ be be the deviations in state, input, and output:

$$\begin{aligned}\boldsymbol{\nu}(t) &= \boldsymbol{\xi}(t) - \boldsymbol{\xi}_0(t),\\ \boldsymbol{\omega}(t) &= \boldsymbol{\mu}(t) - \boldsymbol{\mu}_0(t),\\ \boldsymbol{\gamma}(t) &= \boldsymbol{\eta}(t) - \boldsymbol{\eta}_0(t).\end{aligned}$$

  These are assumed "small."

Linearisation of continuous-time state space systems

- We will be doing a naïve Taylor expansion, and to do so requires taking derivatives. We establish some notation for this.
- First denote by

$$\begin{aligned}\boldsymbol{f}_t \colon X \times U &\to \mathbb{R}^n & \boldsymbol{h}_t \colon X \times U &\to \mathbb{R}^n\\ (\boldsymbol{x}, \boldsymbol{u}) &\mapsto \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) & (\boldsymbol{x}, \boldsymbol{u}) &\mapsto \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{u})\end{aligned}$$

  the mappings for fixed $t \in \mathbb{T}$.
- Assumptions:
  1. U is open (X is always assumed to be open) and
  2. $\boldsymbol{f}_t$ and $\boldsymbol{h}_t$ are continuously differentiable for $t \in \mathbb{T}$.

Linearisation of continuous-time state space systems

- Then denote by $\boldsymbol{D}_1\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}) \in \mathsf{L}(\mathbb{R}^n;\mathbb{R}^n)$ the Jacobian with respect to $\boldsymbol{x}$.
- Explicitly if

$$\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}) = (f_{t,1}(\boldsymbol{x},\boldsymbol{u}),\dots,f_{t,n}(\boldsymbol{x},\boldsymbol{u})),x$$

then

$$\boldsymbol{D}_1\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}) = \begin{bmatrix} \frac{\partial f_{t,1}}{\partial x_1}(\boldsymbol{x},\boldsymbol{u}) & \cdots & \frac{\partial f_{t,1}}{\partial x_n}(\boldsymbol{x},\boldsymbol{u}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{t,n}}{\partial x_1}(\boldsymbol{x},\boldsymbol{u}) & \cdots & \frac{\partial f_{t,n}}{\partial x_n}(\boldsymbol{x},\boldsymbol{u}) \end{bmatrix}.$$

- Also denote by $\boldsymbol{D}_2\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}) \in \mathsf{L}(\mathbb{R}^m;\mathbb{R}^n)$ the Jacobian with respect to $\boldsymbol{u}$.
- Explicitly,

$$\boldsymbol{D}_2\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}) = \begin{bmatrix} \frac{\partial f_{t,1}}{\partial u_1}(\boldsymbol{x},\boldsymbol{u}) & \cdots & \frac{\partial f_{t,1}}{\partial u_m}(\boldsymbol{x},\boldsymbol{u}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{t,n}}{\partial u_1}(\boldsymbol{x},\boldsymbol{u}) & \cdots & \frac{\partial f_{t,n}}{\partial u_m}(\boldsymbol{x},\boldsymbol{u}) \end{bmatrix}.$$

Linearisation of continuous-time state space systems

- Similarly, we have the Jacobians of $\boldsymbol{h}$ with respect to $\boldsymbol{x}$ and $\boldsymbol{u}$:

$$\boldsymbol{D}_1\boldsymbol{h}_t(\boldsymbol{x},\boldsymbol{u}), \quad \boldsymbol{D}_2\boldsymbol{h}_t(\boldsymbol{x},\boldsymbol{u}).$$

- Now we get rid of the pesky t subscripts that we used to make the derivatives clearer. That is, we denote

$$\begin{aligned} \boldsymbol{D}_1\boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{u}) &= \boldsymbol{D}_1\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}), \\ \boldsymbol{D}_2\boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{u}) &= \boldsymbol{D}_2\boldsymbol{f}_t(\boldsymbol{x},\boldsymbol{u}), \\ \boldsymbol{D}_1\boldsymbol{h}(t,\boldsymbol{x},\boldsymbol{u}) &= \boldsymbol{D}_1\boldsymbol{h}_t(\boldsymbol{x},\boldsymbol{u}), \\ \boldsymbol{D}_2\boldsymbol{h}(t,\boldsymbol{x},\boldsymbol{u}) &= \boldsymbol{D}_2\boldsymbol{h}_t(\boldsymbol{x},\boldsymbol{u}). \end{aligned}$$

Linearisation of continuous-time state space systems

- Now we just write the equations for $(\boldsymbol{\xi}, \boldsymbol{\mu})$ and $(\gamma, \boldsymbol{\mu})$ and nastily Taylor expand.
- The writing down part is

$$\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$
$$\boldsymbol{\eta}(t) = \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),$$

or

$$\begin{aligned} \dot{\boldsymbol{\xi}}_0(t) + \dot{\boldsymbol{\nu}}(t) &= \boldsymbol{f}(t, \boldsymbol{\xi}_0(t) + \boldsymbol{\nu}(t), \boldsymbol{\mu}_0(t)\boldsymbol{\omega}(t)), \\ \boldsymbol{\eta}_0(t) + \gamma(t) &= \boldsymbol{h}(t, \boldsymbol{\xi}_0(t) + \boldsymbol{\nu}(t), \boldsymbol{\mu}_0(t) + \boldsymbol{\omega}(t)). \end{aligned} \quad (5)$$

- Now we perform the aforementioned nasty Taylor expansion of (6) about $(\boldsymbol{\xi}(t), \boldsymbol{\mu}_0(t))$, omitting all terms of second-order or more:

$$\begin{aligned} \dot{\boldsymbol{\xi}}_0(t) + \dot{\boldsymbol{\nu}}(t) = &\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) + \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\nu}(t) \\ &+ \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\omega}(t), \\ \boldsymbol{\eta}_0(t) + \gamma(t) = &\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) + \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\nu}(t) \\ &\boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\omega}(t). \end{aligned}$$

Linearisation of continuous-time state space systems

- The zeroth-order terms cancel since $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) \in \mathsf{Ctraj}(\Sigma)$ and $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0) \in \mathsf{Cout}(\Sigma)$.
- Finally, we end up with the "linearised equations:"

$$\begin{aligned} \dot{\boldsymbol{\nu}}(t) &= \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\nu}(t) + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\omega}(t), \\ \gamma(t) &= \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\nu}(t) + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\omega}(t). \end{aligned}$$

- We note that the state for the linearised system is $\boldsymbol{\nu}$, the input is ω, and the output is $\boldsymbol{\eta}$.
- Let us denote for brevity

$$\begin{aligned} \boldsymbol{A}_{(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^n; \mathbb{R}^n) & \boldsymbol{C}_{(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^n; \mathbb{R}^k) \\ t &\mapsto \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)), & t &\mapsto \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)), \\ \boldsymbol{B}_{(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^m; \mathbb{R}^n) & \boldsymbol{D}_{(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^m; \mathbb{R}^k) \\ t &\mapsto \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)), & t &\mapsto \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \end{aligned}$$

Linearisation of continuous-time state space systems

Definition

The ***linearisation*** of $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ about $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) \in \text{Ctraj}(\Sigma)$ with $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0) \in \text{Cout}(\Sigma)$ the controlled output is

$$\Sigma_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)} = (\mathbb{R}^n, \mathbb{R}^m, \mathbb{T}', \mathsf{L}^\infty_{\text{loc}}(\mathbb{T}'; \mathbb{R}^m), \boldsymbol{f}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}, \boldsymbol{h}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}).$$

where

$$\boldsymbol{f}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t, \boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{A}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{v} + \boldsymbol{B}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{w},$$
$$\boldsymbol{h}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t, \boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{C}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{v} + \boldsymbol{D}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{w}.$$

- Thus a controlled trajectory $(\boldsymbol{\nu}, \boldsymbol{\omega}) \in \text{Ctraj}(\Sigma_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)})$ with corresponding controlled output $(\gamma, \boldsymbol{\mu})$ satisfies the equations

$$\dot{\boldsymbol{\nu}}(t) = \boldsymbol{A}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\nu}(t) + \boldsymbol{B}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\omega}(t),$$
$$\gamma(t) = \boldsymbol{C}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\nu}(t) + \boldsymbol{D}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\omega}(t).$$

- This is a time-varying linear system, an object we will study in detail later.

Reading for Lecture 22

1. Section V-6.5.1.1.

Lecture 22 supplement
Aside on meaning of linearisation with respect to state

- Note that in the sloppy derivation of the linearisation, we supposed that the perturbations $\boldsymbol{\nu}$ and $\boldsymbol{\omega}$ were small, but that in the definition of linearisation, the state space where $\boldsymbol{\nu}$ lives is $\mathbb{R}^n$ and the input set where $\boldsymbol{\omega}$ lives is $\mathbb{R}^m$.
- To understand this bit of weirdness, and for other reasons, one should think for a moment about what linearisation means. We shall do this for the part of the linearisation that corresponds to the state.
- We suppose that we have a reference trajectory $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)$ and we let $(t_0, \boldsymbol{x}_0) \in \mathbb{T} \times X$ be such that

$$\boldsymbol{\xi}(t) = \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0, \boldsymbol{\mu}_0).$$

- Let $\boldsymbol{v} \in \mathbb{R}^n$ and consider a variation of the initial condition by

$$s \mapsto \boldsymbol{x}_0 + s\boldsymbol{v}.$$

- Thus we shall vary the initial condition from $\boldsymbol{x}_0$ in the direction of $\boldsymbol{v}$.

Aside on meaning of linearisation with respect to state

- Now, for fixed s small, define

$$\boldsymbol{\xi}_s(t) = \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0 + s\boldsymbol{x}, \boldsymbol{\mu}_0).$$

  This is a new controlled trajectory $(\boldsymbol{\xi}_s, \boldsymbol{\mu}_0)$ with an initial condition close to that of $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)$.
- We should think of $s \mapsto \boldsymbol{\xi}_s$ as being a family of trajectories indexed by s. For each s, we should think of $\boldsymbol{\xi}_s$ as a curve in X.
- We think of

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}_s(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0 + s\boldsymbol{v}, \boldsymbol{\mu}_0)$$

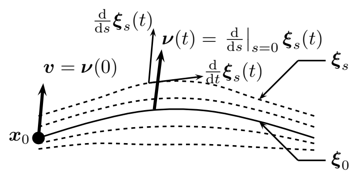
  as being the derivative along the trajectory and

$$\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\xi}_s(t) = \frac{\mathrm{d}}{\mathrm{d}s}\Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0 + s\boldsymbol{v}, \boldsymbol{\mu}_0)$$

  as being the derivative in the direction of the variation of the trajectory.

Aside on meaning of linearisation with respect to state

- One can then show, with some effort, that the solution to the initial value problem
$$\dot{\boldsymbol{\nu}}(t) = \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{\xi}_0(t), \boldsymbol{\mu}_0(t)) \cdot \boldsymbol{\nu}(t), \quad \boldsymbol{\nu}(0) = \boldsymbol{v},$$
is exactly
$$\boldsymbol{\nu}(t) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi^{\Sigma}(t, t_0, \boldsymbol{x}_0 + s\boldsymbol{v}, \boldsymbol{\mu}_0).$$
- Thus $\boldsymbol{\nu}(t)$ measures how the trajectory varies at time t when the initial condition varies in the direction $\boldsymbol{v}$ at time t_0.
- A picture gives all the information.



Reading for Lecture 22 supplement

1. Section V-5.1.1.3.

Lecture 23
Linearisation of continuous-time state space systems (cont'd)

- We note that we can combine the equations for a controlled output $(\boldsymbol{\eta}, \boldsymbol{\mu})$ with the equations for linearisation about $(\boldsymbol{\eta}, \boldsymbol{\mu})$:

$$\begin{aligned}
\dot{\boldsymbol{\xi}}(t) &= \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)), \\
\dot{\boldsymbol{\nu}}(t) &= \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\nu}(t) + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\omega}(t), \\
\boldsymbol{\eta}(t) &= \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)), \\
\boldsymbol{\gamma}(t) &= \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\nu}(t) + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\omega}(t).
\end{aligned}$$

- The first two equations should be thought of as ordinary differential equations for the state $t \mapsto (\xi(t), \boldsymbol{\nu}(t))$ and the second two equations should be thought of as the output equations for the output $t \mapsto (\boldsymbol{\eta}(t), \boldsymbol{\gamma}(t))$.

Linearisation of continuous-time state space systems

- This suggests a linearisation, not about a reference trajectory, but of the whole system, with twice as many state and twice as many outputs.

Definition

The ***linearisation*** of Σ is

$$\Sigma_L = (X \times \mathbb{R}^n, U \times \mathbb{R}^m, \mathbb{T}, \mathscr{U} \times \mathsf{L}^\infty_{\text{loc}}((\mathbb{T}); \mathbb{R}^m), \boldsymbol{f}_L, \boldsymbol{h}_L),$$

where

$$\begin{aligned}
\boldsymbol{f}_L \colon & \mathbb{T} \times (X \times \mathbb{R}^n) \times (U \times \mathbb{R}^m) \to \mathbb{R}^n \times \mathbb{R}^n \\
& (t, (\boldsymbol{x}, \boldsymbol{v}), (\boldsymbol{u}, \boldsymbol{w})) \mapsto (\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}), \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{w}), \\
\boldsymbol{h}_L \colon & \mathbb{T} \times (X \times \mathbb{R}^n) \times (U \times \mathbb{R}^m) \to \mathbb{R}^k \times \mathbb{R}^k \\
& (t, (\boldsymbol{x}, \boldsymbol{v}), (\boldsymbol{u}, \boldsymbol{w})) \mapsto (\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}), \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{w}).
\end{aligned}$$

Linearisation of continuous-time state space systems

- A controlled trajectory $((\boldsymbol{\xi}, \boldsymbol{\nu}), (\boldsymbol{\mu}, \boldsymbol{\omega}))$ giving rise to a controlled output $((\boldsymbol{\eta}, \gamma), (\boldsymbol{\mu}, \boldsymbol{\omega}))$ should be thought of as
  1. first giving a controlled trajectory $(\boldsymbol{\xi}, \boldsymbol{\mu}) \in \mathrm{Ctraj}(\Sigma)$ which gives the controlled output $(\boldsymbol{\eta}, \boldsymbol{\mu}) \in \mathrm{Cout}(\Sigma)$ and
  2. second, determining the linearization about $(\boldsymbol{\eta}, \boldsymbol{\mu})$ in the previous lecture, when linearising about a reference trajectory.
- Thus, Σ_L contains all of the information contained in $\Sigma_{L,(\boldsymbol{\xi},\boldsymbol{\mu})}$, and as well contains the equations for the reference trajectory itself.
- Next we consider linearisation about equilibria.

Definition

A ***controlled equilibrium*** for a continuous-time state space system $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ is a pair $(\boldsymbol{x}_0, \boldsymbol{u}_0) \in X \times U$ such that $\boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) = \boldsymbol{0}$ for all $t \in \mathbb{T}$.

Linearisation of continuous-time state space systems

- For $(\boldsymbol{x}_0, \boldsymbol{u}_0) \in X \times U$, let us denote by

$$\boldsymbol{\xi}_{\boldsymbol{x}_0}(t) = \boldsymbol{x}_0, \quad \boldsymbol{\mu}_{\boldsymbol{u}_0}(t) = \boldsymbol{u}_0$$

  the constant trajectory and the constant input.

Lemma

$(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is a controlled equilibrium for Σ if and only if $(\boldsymbol{\xi}_{\boldsymbol{x}_0}, \boldsymbol{\mu}_{\boldsymbol{u}_0}) \in \mathrm{Ctraj}(\Sigma)$.

Proof.

Suppose that $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is a controlled equilibrium. Then

$$\dot{\boldsymbol{\xi}}_{\boldsymbol{x}_0}(t) = \boldsymbol{0} = \boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) = \boldsymbol{f}(t, \boldsymbol{\xi}_{\boldsymbol{x}_0}(t), \boldsymbol{\mu}_{\boldsymbol{u}_0}(t)),$$

whence $(\boldsymbol{\xi}_{\boldsymbol{x}_0}, \boldsymbol{\mu}_{\boldsymbol{u}_0}) \in \mathrm{Ctraj}(\Sigma)$.
Now assume that $(\boldsymbol{\xi}_{\boldsymbol{x}_0}, \boldsymbol{\mu}_{\boldsymbol{u}_0}) \in \mathrm{Ctraj}(\Sigma)$. Then

$$\boldsymbol{0} = \dot{\boldsymbol{\xi}}_{\boldsymbol{x}_0}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}_{\boldsymbol{x}_0}(t), \boldsymbol{\mu}_{\boldsymbol{u}_0}(t)) = \boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0),$$

whence $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is a controlled equilibrium. □

Linearisation of continuous-time state space systems

- Thus we can linearise about $t \mapsto (\boldsymbol{x}_0, \boldsymbol{u}_0)$ since it is a controlled trajectory.

Definition

Let $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ be a controlled equilibrium for a continuous-time state space system $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$. The ***linearisation*** of Σ about $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is

$$\Sigma_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)} = (\mathbb{R}^n, \mathbb{R}^m, \mathbb{T}, \mathsf{L}^\infty_{\mathsf{loc}}(\mathbb{T}; \mathbb{R}^m), \boldsymbol{f}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}, \boldsymbol{h}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}),$$

where

$$\begin{aligned}\boldsymbol{f}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}(t, \boldsymbol{v}, \boldsymbol{w}) &= \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{w},\\ \boldsymbol{h}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}(t, \boldsymbol{v}, \boldsymbol{w}) &= \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{w}.\end{aligned}$$

Linearisation of continuous-time state space systems

- For autonomous systems, we can define constant matrices

$$\begin{aligned}\boldsymbol{A}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_1\boldsymbol{f}(\boldsymbol{x}_0, \boldsymbol{u}_0),\\ \boldsymbol{B}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_2\boldsymbol{f}(\boldsymbol{x}_0, \boldsymbol{u}_0),\\ \boldsymbol{C}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_1\boldsymbol{h}(\boldsymbol{x}_0, \boldsymbol{u}_0),\\ \boldsymbol{D}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_2\boldsymbol{h}(\boldsymbol{x}_0, \boldsymbol{u}_0)\end{aligned}$$

- Then controlled outputs for the linearisation satisfy

$$\begin{aligned}\dot{\boldsymbol{\nu}}(t) &= \boldsymbol{A}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\nu}(t) + \boldsymbol{B}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\omega}(t),\\ \gamma(t) &= \boldsymbol{C}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\nu}(t) + \boldsymbol{D}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\omega}(t).\end{aligned}$$

  This is a linear time-invariant system, about which we will have much to say down the road.

Reading for Lecture 23

1. Sections V-6.5.1.1 and V-6.5.1.2.

Lecture 24
Linearisation of continuous-time input/output systems

- It is possible, but difficult for this level of course, to talk about linearisation of input/output systems.
- We shall not do so.
- So. . . you're welcome!

Linearisation of discrete-time state space systems

- This goes rather like for continuous-time systems, of course. But let's go through it.
- Let $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ be a discrete-time state space system and let $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) \in$ Ctraj(Σ) with $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0) \in$ Cout(Σ) the corresponding controlled output. Let $\mathrm{dom}(\boldsymbol{\mu}_0) = \mathbb{T}'$.
- Call $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)$ the ***reference trajectory*** and $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0)$ the ***reference output***.
- We will linearise about this the reference. This means that we consider a controlled trajectory $(\boldsymbol{\xi}, \boldsymbol{\mu})$ that is "nearby" $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0)$. We let $\boldsymbol{\nu}$, $\boldsymbol{\omega}$, and γ be be the deviations in state, input, and output:

$$\begin{aligned}\boldsymbol{\nu}(t) &= \boldsymbol{\xi}(t) - \boldsymbol{\xi}_0(t),\\ \boldsymbol{\omega}(t) &= \boldsymbol{\mu}(t) - \boldsymbol{\mu}_0(t),\\ \gamma(t) &= \boldsymbol{\eta}(t) - \boldsymbol{\eta}_0(t).\end{aligned}$$

  These are assumed "small."

Linearisation of discrete-time state space systems

- Assumptions:
  1. U is open (X is always assumed to be open) and
  2. $(\boldsymbol{x}, \boldsymbol{u}) \mapsto \boldsymbol{f}(t, \boldsymbol{u}, \boldsymbol{u}$ and $(\boldsymbol{x}, \boldsymbol{u}) \mapsto \boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u})$ are continuously differentiable for $t \in \mathbb{T}$.
- Just as in the continuous-time case, we have the Jacobians of $\boldsymbol{f}$ and $\boldsymbol{h}$ with respect to $\boldsymbol{x}$ and $\boldsymbol{u}$:

$$\begin{aligned}&\boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}),\\ &\boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}),\\ &\boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}),\\ &\boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}).\end{aligned}$$

Linearisation of discrete-time state space systems

- Now we just write the equations for (ξ, μ) and (γ, μ) and nastily Taylor expand.
- The writing down part is

$$\begin{aligned}\xi(t+\Delta) &= f(t, \xi(t), \mu(t)),\\ \eta(t) &= h(t, \xi(t), \mu(t)),\end{aligned}$$

or

$$\begin{aligned}\xi_0(t+\Delta) + \nu(t+\Delta) &= f(t, \xi_0(t) + \nu(t), \mu_0(t)\omega(t)),\\ \eta_0(t) + \gamma(t) &= h(t, \xi_0(t) + \nu(t), \mu_0(t) + \omega(t)).\end{aligned} \quad (6)$$

- Now we perform the aforementioned nasty Taylor expansion of (6) about $(\xi(t), \mu_0(t))$, omitting all terms of second-order or more:

$$\begin{aligned}\xi_0(t+\Delta) + \nu(t+\Delta) &= f(t, \xi_0(t), \mu_0(t)) + D_1 f(t, \xi_0(t), \mu_0(t)) \cdot \nu(t)\\ &\quad + D_2 f(t, \xi_0(t), \mu_0(t)) \cdot \omega(t),\\ \eta_0(t) + \gamma(t) &= h(t, \xi_0(t), \mu_0(t)) + D_1 h(t, \xi_0(t), \mu_0(t)) \cdot \nu(t)\\ &\quad D_2 h(t, \xi_0(t), \mu_0(t)) \cdot \omega(t).\end{aligned}$$

Linearisation of discrete-time state space systems

- The zeroth-order terms cancel since $(\xi_0, \mu_0) \in \mathsf{Ctraj}(\Sigma)$ and $(\eta_0, \mu_0) \in \mathsf{Cout}(\Sigma)$.
- Finally, we end up with the "linearised equations:"

$$\begin{aligned}\nu(t+\Delta) &= D_1 f(t, \xi_0(t), \mu_0(t)) \cdot \nu(t) + D_2 f(t, \xi_0(t), \mu_0(t)) \cdot \omega(t),\\ \gamma(t) &= D_1 h(t, \xi_0(t), \mu_0(t)) \cdot \nu(t) + D_2 h(t, \xi_0(t), \mu_0(t)) \cdot \omega(t).\end{aligned}$$

- We note that the state for the linearised system is ν, the input is ω, and the output is η.
- Let us denote for brevity

$$\begin{aligned}A_{(\xi_0,\mu_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^n; \mathbb{R}^n) & C_{(\xi_0,\mu_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^n; \mathbb{R}^k)\\ t &\mapsto D_1 f(t, \xi_0(t), \mu_0(t)), & t &\mapsto D_1 h(t, \xi_0(t), \mu_0(t)),\\ B_{(\xi_0,\mu_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^m; \mathbb{R}^n) & D_{(\xi_0,\mu_0)} \colon \mathbb{T}' &\to \mathsf{L}(\mathbb{R}^m; \mathbb{R}^k)\\ t &\mapsto D_2 f(t, \xi_0(t), \mu_0(t)), & t &\mapsto D_2 h(t, \xi_0(t), \mu_0(t))\end{aligned}$$

Linearisation of discrete-time state space systems

Definition

The ***linearisation*** of $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ about $(\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) \in \mathrm{Ctraj}(\Sigma)$ with $(\boldsymbol{\eta}_0, \boldsymbol{\mu}_0) \in \mathrm{Cout}(\Sigma)$ the controlled output is

$$\Sigma_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)} = (\mathbb{R}^n, \mathbb{R}^m, \mathbb{T}', \ell_{\mathrm{loc}}(\mathbb{T}';\mathbb{R}^m), \boldsymbol{f}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}, \boldsymbol{h}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}).$$

where

$$\begin{aligned}\boldsymbol{f}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t,\boldsymbol{v},\boldsymbol{w}) &= \boldsymbol{A}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{v} + \boldsymbol{B}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{w},\\ \boldsymbol{h}_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t,\boldsymbol{v},\boldsymbol{w}) &= \boldsymbol{C}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{v} + \boldsymbol{D}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{w}.\end{aligned}$$

- Thus a controlled trajectory $(\boldsymbol{\nu}, \boldsymbol{\omega}) \in \mathrm{Ctraj}(\Sigma_{L,(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)})$ with corresponding controlled output $(\boldsymbol{\gamma}, \boldsymbol{\mu})$ satisfies the equations

$$\begin{aligned}\boldsymbol{\nu}(t+\Delta) &= \boldsymbol{A}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\nu}(t) + \boldsymbol{B}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\omega}(t),\\ \boldsymbol{\gamma}(t) &= \boldsymbol{C}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\nu}(t) + \boldsymbol{D}_{(\boldsymbol{\xi}_0,\boldsymbol{\mu}_0)}(t)\boldsymbol{\omega}(t).\end{aligned}$$

- This is a time-varying linear system, an object we will study in detail later.

Linearisation of discrete-time state space systems

- We note that we can combine the equations for a controlled output $(\boldsymbol{\eta}, \boldsymbol{\mu})$ with the equations for linearisation about $(\boldsymbol{\eta}, \boldsymbol{\mu})$:

$$\begin{aligned}\boldsymbol{\xi}(t+\Delta) &= \boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),\\ \boldsymbol{\nu}(t+\Delta) &= \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\nu}(t) + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\omega}(t),\\ \boldsymbol{\eta}(t) &= \boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)),\\ \boldsymbol{\gamma}(t) &= \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\nu}(t) + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{\xi}(t), \boldsymbol{\mu}(t)) \cdot \boldsymbol{\omega}(t).\end{aligned}$$

- The first two equations should be thought of as ordinary difference equations for the state $t \mapsto (\xi(t), \boldsymbol{\nu}(t))$ and the second two equations should be thought of as the output equations for the output $t \mapsto (\boldsymbol{\eta}(t), \boldsymbol{\gamma}(t))$.

Linearisation of discrete-time state space systems

- This suggests a linearisation, not about a reference trajectory, but of the whole system, with twice as many state and twice as many outputs.

Definition

The ***linearisation*** of Σ is

$$\Sigma_L = (X \times \mathbb{R}^n, U \times \mathbb{R}^m, \mathbb{T}, \mathscr{U} \times \ell_{\mathrm{loc}}((\mathbb{T}); \mathbb{R}^m), \boldsymbol{f}_L, \boldsymbol{h}_L),$$

where

$$\begin{aligned} \boldsymbol{f}_L \colon & \mathbb{T} \times (X \times \mathbb{R}^n) \times (U \times \mathbb{R}^m) \to \mathbb{R}^n \times \mathbb{R}^n \\ & (t, (\boldsymbol{x}, \boldsymbol{v}), (\boldsymbol{u}, \boldsymbol{w})) \mapsto (\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}), \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{w}), \\ \boldsymbol{h}_L \colon & \mathbb{T} \times (X \times \mathbb{R}^n) \times (U \times \mathbb{R}^m) \to \mathbb{R}^k \times \mathbb{R}^k \\ & (t, (\boldsymbol{x}, \boldsymbol{v}), (\boldsymbol{u}, \boldsymbol{w})) \mapsto (\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}), \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{x}, \boldsymbol{u}) \cdot \boldsymbol{w}). \end{aligned}$$

Linearisation of discrete-time state space systems

- A controlled trajectory $((\boldsymbol{\xi}, \boldsymbol{\nu}), (\boldsymbol{\mu}, \boldsymbol{\omega}))$ giving rise to a controlled output $((\boldsymbol{\eta}, \boldsymbol{\gamma}), (\boldsymbol{\mu}, \boldsymbol{\omega}))$ should be thought of as
  1. first giving a controlled trajectory $(\boldsymbol{\xi}, \boldsymbol{\mu}) \in \mathrm{Ctraj}(\Sigma)$ which gives the controlled output $(\boldsymbol{\eta}, \boldsymbol{\mu}) \in \mathrm{Cout}(\Sigma)$ and
  2. second, determining the linearization about $(\boldsymbol{\eta}, \boldsymbol{\mu})$ in the previous lecture, when linearising about a reference trajectory.
- Thus, Σ_L contains all of the information contained in $\Sigma_{L,(\boldsymbol{\xi},\boldsymbol{\mu})}$, and as well contains the equations for the reference trajectory itself.
- Next we consider linearisation about equilibria.

Definition

A ***controlled equilibrium*** for a discrete-time state space system $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$ is a pair $(\boldsymbol{x}_0, \boldsymbol{u}_0) \in X \times U$ such that $\boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) = \boldsymbol{x}_0$ for all $t \in \mathbb{T}$.

Linearisation of discrete-time state space systems

- For $(\boldsymbol{x}_0, \boldsymbol{u}_0) \in X \times U$, let us denote by

$$\xi_{\boldsymbol{x}_0}(t) = \boldsymbol{x}_0, \quad \mu_{\boldsymbol{u}_0}(t) = \boldsymbol{u}_0$$

the constant trajectory and the constant input.

Lemma

$(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is a controlled equilibrium for Σ if and only if $(\xi_{\boldsymbol{x}_0}, \mu_{\boldsymbol{u}_0}) \in \mathrm{Ctraj}(\Sigma)$.

Proof.

Suppose that $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is a controlled equilibrium. Then

$$\xi_{\boldsymbol{x}_0}(t + \Delta) = \boldsymbol{x}_0 = \boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) = \boldsymbol{f}(t, \xi_{\boldsymbol{x}_0}(t), \mu_{\boldsymbol{u}_0}(t)),$$

whence $(\xi_{\boldsymbol{x}_0}, \mu_{\boldsymbol{u}_0}) \in \mathrm{Ctraj}(\Sigma)$.
Now assume that $(\xi_{\boldsymbol{x}_0}, \mu_{\boldsymbol{u}_0}) \in \mathrm{Ctraj}(\Sigma)$. Then

$$\boldsymbol{x}_0 = \xi_{\boldsymbol{x}_0}(t + \Delta) = \boldsymbol{f}(t, \xi_{\boldsymbol{x}_0}(t), \mu_{\boldsymbol{u}_0}(t)) = \boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0),$$

whence $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is a controlled equilibrium. □

Linearisation of discrete-time state space systems

- Thus we can linearise about $t \mapsto (\boldsymbol{x}_0, \boldsymbol{u}_0)$ since it is a controlled trajectory.

Definition

Let $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ be a controlled equilibrium for a continuous-time state space system $\Sigma = (X, U, \mathbb{T}, \mathscr{U}, \boldsymbol{f}, \boldsymbol{h})$. The ***linearisation*** of Σ about $(\boldsymbol{x}_0, \boldsymbol{u}_0)$ is

$$\Sigma_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)} = (\mathbb{R}^n, \mathbb{R}^m, \mathbb{T}, \ell_{\mathrm{loc}}(\mathbb{T}; \mathbb{R}^m), \boldsymbol{f}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}, \boldsymbol{h}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}),$$

where

$$\begin{aligned} \boldsymbol{f}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}(t, \boldsymbol{v}, \boldsymbol{w}) &= \boldsymbol{D}_1\boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{w}, \\ \boldsymbol{h}_{L,(\boldsymbol{x}_0,\boldsymbol{u}_0)}(t, \boldsymbol{v}, \boldsymbol{w}) &= \boldsymbol{D}_1\boldsymbol{h}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{v} + \boldsymbol{D}_2\boldsymbol{h}(t, \boldsymbol{x}_0, \boldsymbol{u}_0) \cdot \boldsymbol{w}. \end{aligned}$$

Linearisation of discrete-time state space systems

- For autonomous systems, we can define constant matrices

$$\begin{aligned} \boldsymbol{A}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_1\boldsymbol{f}(\boldsymbol{x}_0,\boldsymbol{u}_0), \\ \boldsymbol{B}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_2\boldsymbol{f}(\boldsymbol{x}_0,\boldsymbol{u}_0), \\ \boldsymbol{C}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_1\boldsymbol{h}(\boldsymbol{x}_0,\boldsymbol{u}_0), \\ \boldsymbol{D}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)} &= \boldsymbol{D}_2\boldsymbol{h}(\boldsymbol{x}_0,\boldsymbol{u}_0) \end{aligned}$$

- Then controlled outputs for the linearisation satisfy

$$\begin{aligned} \boldsymbol{\nu}(t+\Delta) &= \boldsymbol{A}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\nu}(t) + \boldsymbol{B}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\omega}(t), \\ \gamma(t) &= \boldsymbol{C}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\nu}(t) + \boldsymbol{D}_{(\boldsymbol{x}_0,\boldsymbol{u}_0)}\boldsymbol{\omega}(t). \end{aligned}$$

  This is a linear time-invariant system, about which we will have much to say down the road.

Linearisation of discrete-time input/output systems

- It is possible, but difficult for this level of course, to talk about linearisation of input/output systems.
- We shall not do so.
- So... you're welcome! $(\times 2)$

Reading for Lecture 24

1. Sections V-6.5.3.1 and V-6.5.3.2.

Lecture 25
Linear systems

- We have considered four classes of systems:
  1. continuous-time state space systems;
  2. continuous-time input/output systems;
  3. discrete-time state space systems;
  4. discrete-time input/output systems.
- For these quite general classes of systems, all we could really say were very general things, e.g., about causality, stationarity, continuity, etc.
- We shall now turn our attention exclusively to linear systems, where we will be able to say more about more useful attributes for systems.
- When talking about linear systems, we shall suppose that state spaces, input sets, and output sets are finite-dimensional $\mathbb{R}$-vector spaces. You are welcome to think of these as being $\mathbb{R}^n$, $\mathbb{R}^m$, and $\mathbb{R}^k$, respectively.

Linear continuous-time state space systems

- We start with continuous-time state space systems. We might say that such a system is "linear" when it depends linearly on state and control.

Definition

A ***linear continuous-time state space system*** is

$$\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathbb{T}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}),$$

where

(i) X is a finite-dimensional $\mathbb{R}$-vector space (***state space***),

(ii) U is a finite-dimensional $\mathbb{R}$-vector space (***input set***),

(iii) Y is a finite-dimensional $\mathbb{R}$-vector space (***output set***),

(iv) $\mathbb{T} \subseteq \mathbb{R}$ is a continuous time-domain (***time-domain***),

(v) $\mathscr{U} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ (***inputs***),

(vi) $\mathsf{A} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{X}))$,

(vii) $\mathsf{B} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{X}))$,

(viii) $\mathsf{C} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{Y}))$, and

(ix) $\mathsf{D} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$.

Linear continuous-time state space systems

- A linear continuous-time state space system is an instance of a continuous-time state space system by defining the dynamics and the output map by

$$f(t, x, u) = \mathsf{A}(t)x + \mathsf{B}(t)u, \quad h(t, x, u) = \mathsf{C}(t)x + \mathsf{D}(t)u.$$

  respectively.
- Note that, for linear systems, we do not fuss with partially defined inputs, state, and outputs, although one *can* do this. The reason that this is possible (whereas it is simply not possible for systems that are not linear) is that solutions for linear differential equations, homogeneous or inhomogeneous, cannot blow up in finite time. This is pointed out
  1. for homogeneous equations in item 1 on Slide 33 and
  2. for inhomogeneous equations by virtue of the variation of constants formula on Slide 37.

Linear continuous-time state space systems

- We see, then, that $(\xi, \mu) \in \mathsf{Ctraj}(\Sigma)$ if and only if

$$\dot{\xi}(t) = \mathsf{A}(t) \cdot \xi(t) + \mathsf{B}(t) \cdot \mu(t).$$

- The corresponding controlled output $(\eta, \mu) \in \mathsf{Cout}(\Sigma)$ is

$$\eta(t) = \mathsf{C}(t) \cdot \xi(t) + \mathsf{D}(t) \cdot \mu(t).$$

- Existence and uniqueness of controlled trajectories is established by the following comparatively simple result.

Theorem

Let Σ be a linear continuous-time state space system, let $t_0 \in \mathbb{T}$, let $x_0 \in \mathsf{X}$, and suppose that B and $\mathscr{U}$ are such that $t \in \mathsf{B}(t) \cdot \mu(t)$ is locally integrable for every $\mu \in \mathscr{U}$. Then, for $\mu \in \mathscr{U}$, there exists a unique locally absolutely continuous $\xi \colon \mathbb{T} \to \mathsf{X}$ satisfying

$$\dot{\xi}(t) = \mathsf{A}(t) \cdot \xi(t) + \mathsf{B}(t) \cdot \mu(t), \quad \xi(t_0) = x_0.$$

Linear continuous-time state space systems

- Recall that the homogeneous linear ordinary differential equation

$$\dot{\xi}(t) = \mathsf{A}(t) \cdot \xi(t)$$

has a flow of the form

$$(t, t_0, x_0) \mapsto \Phi^{\mathsf{c}}_{\mathsf{A}}(t, t_0)(x_0),$$

where $\Phi^{\mathsf{c}}_{\mathsf{A}} \colon \mathbb{T} \times \mathbb{T} \to \mathsf{L}(\mathsf{X}; \mathsf{X})$ is the state transition map (item 2 on Slide33).

- Assuming that Σ satisfies the conditions of the preceding theorem, we can then write down an explicit formula for flow using the the variation of constants formula on Slide 37:

$$\Phi^{\Sigma}(t, t_0, x_0, \mu) = \Phi^{\mathsf{c}}_{\mathsf{A}}(t, t_0)(x_0) + \int_{t_0}^{t} \Phi^{\mathsf{c}}_{\mathsf{A}}(t, \tau) \circ \mathsf{B}(\tau) \cdot \mu(\tau) \, \mathrm{d}\tau$$

Linear continuous-time state space systems

- The corresponding output is then

$$\eta(t) = \underbrace{\mathsf{C}(t) \circ \Phi^{\mathsf{c}}_{\mathsf{A}}(t, t_0)(x_0)}_{\textit{term 1}} + \underbrace{\int_{t_0}^{t} \mathsf{C}(t) \circ \Phi^{\mathsf{c}}_{\mathsf{A}}(t, \tau) \circ \mathsf{B}(\tau) \cdot \mu(\tau) \, \mathrm{d}\tau}_{\textit{term 2}} + \underbrace{\mathsf{D}(t) \cdot \mu(t)}_{\textit{term 3}}.$$

- We note that the output is comprised of three bits, each interesting in its own right:
  1. *term 1*: Here is a component of the output determined by the initial condition. Indeed, the other parts of the solution are independent of initial condition.
  2. *term 2*: This is some weird integral. We shall encounter this sort of thing when we consider linear continuous-time input/output systems below.
  3. *term 3*: This term consists of the input at time t directly influencing the output at time t in a memoryless fashion. This is sometimes called a "feedforward" term.

Linear continuous-time state space systems

- Next we consider systems with constant coefficients.

Definition

A ***linear continuous-time state space system with constant coefficients*** is

$$\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathbb{T}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}),$$

where

(i) X is a finite-dimensional $\mathbb{R}$-vector space (***state space***),
(ii) U is a finite-dimensional $\mathbb{R}$-vector space (***input set***),
(iii) Y is a finite-dimensional $\mathbb{R}$-vector space (***output set***),
(iv) $\mathbb{T} \subseteq \mathbb{R}$ is a continuous time-domain (***time-domain***),
(v) $\mathscr{U} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ (***inputs***),
(vi) $\mathsf{A} \in \mathsf{L}(\mathsf{X}; \mathsf{X})$,
(vii) $\mathsf{B} \in \mathsf{L}(\mathsf{U}; \mathsf{X})$,
(viii) $\mathsf{C} \in \mathsf{L}(\mathsf{X}; \mathsf{Y})$, and
(ix) $\mathsf{D} \in \mathsf{L}(\mathsf{U}; \mathsf{Y})$.

Linear continuous-time state space systems

- In short, a "linear continuous time state space system with constant coefficients" is a linear continuous-time state space system where A, B, C, and D are independent of time.
- As such, a linear continuous-time state space system with constant coefficients is also an autonomous continuous-time state space system with
$$f(x,u) = \mathsf{A}(x) + \mathsf{B}(u), \quad h(x,u) = \mathsf{C}(x) + \mathsf{D}(u).$$
- Thus a controlled trajectory $(\xi, \mu) \in \mathrm{Ctraj}(\Sigma)$ with corresponding controlled output $(\eta, \mu) \in \mathrm{Cout}(\Sigma)$ jointly satisfy
$$\dot{\xi}(t) = \mathsf{A} \circ \xi(t) + \mathsf{B} \circ \mu(t),$$
$$\eta(t) = \mathsf{C} \circ \xi(t) + \mathsf{D} \circ \mu(t).$$
- Note that, when A is independent of time,
$$\Phi^{\mathsf{c}}_{\mathsf{A}}(t, t_0) = \mathsf{e}^{\mathsf{A}(t-t_0)}.$$

Linear continuous-time state space systems

- The existence and uniqueness theorem on Slide 200 can then be simplified to just the requirement that $\mathscr{U} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$.
- Thus the formula for the flow and the output equations is as in the non-constant coefficient case, but making this substitution for the state transition map.
- Thus we have
$$\Phi^{\Sigma}(t, t_0, x_0) = \mathsf{e}^{\mathsf{A}(t-t_0)}(x_0) + \int_{t_0}^{t} \mathsf{e}^{\mathsf{A}(t-\tau)} \circ \mathsf{B} \circ \mu(\tau)\, \mathrm{d}\tau$$
and
$$\eta(t) = \mathsf{C} \circ \mathsf{e}^{\mathsf{A}(t-t_0)}(x_0) + \int_{t_0}^{t} \mathsf{C} \circ \mathsf{e}^{\mathsf{A}(t-\tau)} \circ \mathsf{B} \circ \mu(\tau)\, \mathrm{d}\tau + \mathsf{D} \circ \mu(t).$$

Linear continuous-time state space systems

- Because linear continuous-time state space systems with constant coefficients are stationary whenever the input set $\mathscr{U}$ is translation-invariant, one commonly makes two assumptions:
  1. $\mathbb{T} = \mathbb{R}$;
  2. one considers the initial time to be $t_0 = 0$.
- As a consequence of this, one typically works with inputs defined on $\mathbb{R}_{\geq 0}$ and with controlled trajectories and outputs determined by

$$\Phi^{\Sigma}(t, 0, x_0) = \mathsf{e}^{\mathsf{A}t}(x_0) + \int_0^t \mathsf{e}^{\mathsf{A}(t-\tau)} \circ \mathsf{B} \circ \mu(\tau)\, \mathrm{d}\tau$$

and

$$\eta(t) = \underbrace{\mathsf{C} \circ \mathsf{e}^{\mathsf{A}t}(x_0)}_{\textit{term 1}} + \underbrace{\int_0^t \mathsf{C} \circ \mathsf{e}^{\mathsf{A}(t-\tau)} \circ \mathsf{B} \circ \mu(\tau)\, \mathrm{d}\tau}_{\textit{term 2}} + \underbrace{\mathsf{D} \circ \mu(t)}_{\textit{term 3}}.$$

- The three terms have a similar interpretation as in the time-varying case, except that *term 2* is a convolution. We shall care about this.

Reading for Lecture 25

1. Sections V-6.6.1 and V-6.6.2.

Lecture 26
Linear continuous-time state space systems (cont'd)

- For linear continuous-time state space systems, we had the following formulae for outputs.
  1. Time-varying case, input μ, initial condition x_0 at initial time t_0:

$$\eta(t) = \underbrace{\mathsf{C}(t) \circ \Phi_{\mathsf{A}}(t, t_0)(x_0)}_{\textit{term 1}} + \underbrace{\int_{t_0}^{t} \mathsf{C}(t) \circ \Phi_{\mathsf{A}}(t, \tau) \circ \mathsf{B}(\tau) \cdot \mu(\tau)\, \mathrm{d}\tau}_{\textit{term 2}} + \underbrace{\mathsf{D}(t) \cdot \mu(t)}_{\textit{term 3}}.$$

  2. Constant coefficient case, input μ, initial condition x_0 at initial time 0:

$$\eta(t) = \underbrace{\mathsf{C} \circ \mathrm{e}^{\mathsf{A}t}(x_0)}_{\textit{term 1}} + \underbrace{\int_{0}^{t} \mathsf{C} \circ \mathrm{e}^{\mathsf{A}(t-\tau)} \circ \mathsf{B} \circ \mu(\tau)\, \mathrm{d}\tau}_{\textit{term 2}} + \underbrace{\mathsf{D} \circ \mu(t)}_{\textit{term 3}}.$$

- Our objective now is to explore more fully the components labelled in the preceding formulae as "*term 2*."

Linear continuous-time state space systems

Definition

Let σ be an linear continuous-time state space system.

(i) The ***proper impulse transmission map*** is

$$\begin{aligned} \mathrm{pitm}_{\Sigma} \colon \mathbb{T} \times \mathbb{T} &\to \mathsf{L}(\mathsf{U}; \mathsf{Y}) \\ (t, \tau) &\mapsto 1_{\geq 0}(t - \tau)\mathsf{C}(t) \circ \Phi_{\mathsf{A}}(t, \tau) \circ \mathsf{B}(\tau). \end{aligned}$$

(ii) One can define the "impulse transmission map," but we will not. It involves a few too many elementary, but notationally complex, constructions with distributions. We shall make these constructions below in the constant coefficient case, where we can more easily understand what we are doing.

- Obviously, the output for an input μ and initial condition x_0 at t_0 is

$$\eta(t) = \mathsf{C}(t) \circ \Phi_{\mathsf{A}}(t, t_0)(x_0) + \int_{t_0}^{t} \mathrm{pitm}_{\Sigma}(t, \tau) \cdot \mu(\tau) \mathrm{d}\tau + \mathsf{D}(t) \cdot \mu(t). \qquad (7)$$

Linear continuous-time state space systems

- Let us make a few comments about the preceding.
  1. When $\mathsf{D}(t) = 0$ for all t, when $x_0 = 0$, and when the input is the delta-distribution δ_{t_0} at t_0,
  $$\mathsf{pitm}_\Sigma(t, t_0) = \text{“}\int_{t_0}^{t} \mathsf{pitm}_\Sigma(t,\tau)\cdot\tau_{t_0}^{*}\delta(\tau)\,\mathrm{d}\tau\text{”}$$
  gives the output for an input that is an impulse at t_0.
  2. The formula also gives context to the terminology “impulse transmission map.” We see that $\mathsf{pitm}_\Sigma(t,\tau)$ in the integrand serves to “transmit” the effect of the input at time τ to the output at time t.
- Now we look at the impulse transmission map in the constant coefficient case.
- Here we will think a little more carefully about the rôle of distributions, since we have already seen what we need to make sense of this.

Linear continuous-time state space systems

- We outline the facts we need to define the impulse transmission map in the constant coefficient case.
  1. Recall from Slide 78 that the solution (unique, if we restrict to distributions with support bounded on the left) to the distributional ordinary differential equation
  $$\theta_u^{(1)} = \mathsf{A}(\theta_u) + \mathsf{B}(u \otimes \delta)$$
  is the regular distribution associated with the function $t \mapsto 1_{\geq 0}(t)\mathrm{e}^{\mathsf{A}t} \circ \mathsf{B}(u)$. (We also refer to the referred for a discussion of the notation in this equation.)
  2. Thus, when $\mathsf{D} = 0$ and $x_0 = 0$, the output for the input $\mu = u \otimes \delta$ is
  $$\eta(t) = 1_{\geq 0}(t)\mathsf{C} \circ \mathrm{e}^{\mathsf{A}t} \circ \mathsf{B}(u)$$
  3. When $\mathsf{D} \neq 0$, the output has to incorporate the *term 3* in the output formula, which is $\mathsf{D} \circ \mu$.
  4. But the input μ is now a distribution, namely the U-valued distribution $\mu = u \otimes \delta$.
  5. So the output also will be a distribution, namely the Y-valued distribution
  $$\theta_{1_{\geq 0}\exp_{\mathsf{A}}\circ\mathsf{B}(u)} + \mathsf{D}(u \otimes \delta).$$

Linear continuous-time state space systems

6. Thus the output for the input $u \otimes \delta$ consists of two bits, one the regular distribution corresponding to the function $t \mapsto 1_{\geq 0}(t)\mathsf{C} \circ e^{\mathsf{A}t} \circ \mathsf{B}(u)$ and the other the singular distribution $\mathsf{D}(u \otimes \delta)$.

- With this as backdrop, we make the following definition.

Definition

Let Σ be a linear continuous-time state space system with constant coefficients.

(i) The ***proper impulse response*** is

$$\begin{aligned} \mathsf{pir}_\Sigma \colon \mathbb{R} &\to \mathsf{L}(\mathsf{U};\mathsf{Y}) \\ t &\mapsto 1_{\geq 0}(t)\mathsf{C} \circ e^{\mathsf{A}t} \circ \mathsf{B}. \end{aligned}$$

(ii) The ***impulse response*** is the $\mathsf{L}(\mathsf{U};\mathsf{Y})$-valued distribution given by

$$\mathsf{ir}_\Sigma(u) = \theta_{\mathsf{pir}_\Sigma}(u) + \mathsf{D}(u \otimes \delta).$$

Linear continuous-time state space systems

- Note that a linear continuous-time state space system with constant coefficients is, in particular, a linear continuous-time state space system. As such, in the constant coefficient case, we still have the notion of an impulse transmission map. Indeed, we have

$$\mathsf{pitm}_\Sigma(t,\tau) = \mathsf{pir}_\Sigma(t-\tau)$$

in the constant coefficient case.
- Note that the output associated to the input μ with initial condition x_0 at $t = 0$ is

$$\eta(t) = \mathsf{C} \circ e^{\mathsf{A}t}(x_0) + \int_0^t \mathsf{pir}_\Sigma(t-\tau) \circ \mu(\tau)\mathrm{d}\tau + \mathsf{D} \circ \mu(t). \qquad (8)$$

- The middle term is an old friend, namely a convolution!
- Summary:
  1. Both the impulse transmission map (in the time-varying case) and the impulse response (in the constant coefficient case) are the output for an impulse input.
  2. The formulae (7) and (8) illustrate that this response to an impulse forms an integral (pun!) part of the output for a general input.

Linear continuous-time input/output systems

- To use linearity, we need to assume that $\mathscr{U}$ and $\mathscr{Y}$ are such that linearity from $\mathscr{U}$ to $\mathscr{Y}$ makes sense.
- Recall that, if $\mathscr{U} \subseteq \mathsf{L}^1_{\mathrm{loc}}((\mathbb{T}); U)$, then we had denoted, for $\mathbb{S} \subseteq \mathbb{T}$,

$$\mathscr{U}(\mathbb{S}) = \{\mu \in \mathscr{U} \mid \mathrm{dom}(\mu) = \mathbb{S}\}.$$

Definition

An ***linear continuous-time input/output system*** is $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, g)$, where

(i) U is a finite dimensional $\mathbb{R}$-vector space (***input set***),
(ii) Y is a finite dimensional $\mathbb{R}$-vector space (***output set***),
(iii) $\mathbb{T} \subseteq \mathbb{R}$ is a continuous time-domain,
(iv) $\mathscr{U}$ is such that, for every $\mathbb{S} \subseteq \mathbb{T}$, $\mathscr{U}(\mathbb{S})$ is a subspace of $\mathsf{U}^{\mathbb{S}}$,
(v) $\mathscr{Y}$ is such that, for every $\mathbb{S} \subseteq \mathbb{T}$, $\mathscr{Y}(\mathbb{S})$ is a subspace of $\mathsf{Y}^{\mathbb{S}}$, and

Linear continuous-time input/output systems

Definition (cont'd)

(vi) (a) for $\mathbb{S} \subseteq \mathbb{T}$, if $g_{\mathbb{S}} = g|\mathscr{U}(\mathbb{S})$, then $g_{\mathbb{S}}(\mu) \in \mathscr{Y}(\mathbb{S})$,
(b) if $\mathbb{S}' \subseteq \mathbb{S} \subseteq \mathbb{T}$, then

$$g_{\mathbb{S}'}(\mu|\mathbb{S}') = g_{\mathbb{S}}(\mu)|\mathbb{S}',$$

and
(c) for $\mathbb{S} \subseteq \mathbb{T}$, $g_{\mathbb{S}} \colon \mathscr{U}(\mathbb{S}) \to \mathscr{Y}(\mathbb{S})$ is a continuous linear mapping.

- In brief, a linear continuous-time input/output system is a continuous-time input/output system that is. . . linear.

Reading for Lecture 26

1. Section V-6.6.3.
2. Section V-6.7.1.

Lecture 27
Integral kernel systems

- The basic idea:
$$g_{\mathsf{K}}(\mu)(t) = \int_{\mathbb{T}} \mathsf{K}(t,\tau)\mu(\tau)\mathrm{d}\tau.$$
We call K the "integral kernel."
- This should remind you of the proper impulse transmission map for linear continuous-time state space systems.

Definition

Let U, Y be finite-dimensional $\mathbb{R}$-vector spaces, and let $\mathbb{T} \subseteq \mathbb{R}$ be a continuous time-domain.

(i) An ***integral kernel*** is a mapping
$$\mathsf{K}\colon \mathbb{T} \times \mathbb{T} \to \mathsf{L}(\mathsf{U};\mathsf{Y}).$$

(ii) We denote
$$\begin{aligned}\mathsf{K}_t\colon \mathbb{T} &\to \mathsf{L}(\mathsf{U};\mathsf{Y})\\ \tau &\mapsto \mathsf{K}(t,\tau).\end{aligned}$$

Integral kernel systems

Definition (cont'd)

(iii) If $\mathscr{U} \subseteq \mathsf{U}^{\mathbb{T}}$, then K is ***compatible*** with $\mathscr{U}$ if $\tau \mapsto \mathsf{K}_t(\tau)\mu(\tau)$ is in $\mathsf{L}^1(\mathbb{T};\mathsf{Y})$ for every $\mu \in \mathscr{U}$.

(iv) if K is compatible with $\mathscr{U} \subseteq \mathsf{U}^{\mathbb{T}}$, then the ***integral operator*** associated with K is

$$g_{\mathsf{K}} \colon \mathscr{U} \to \mathsf{Y}^{\mathbb{T}}$$

defined by

$$g_{\mathsf{K}}(\mu)(t) = \int_{\mathbb{T}} \mathsf{K}(t,\tau)\mu(\tau)\mathrm{d}\tau.$$

- At this point, we cannot quite call this a linear continuous-time input/output system since we do not have linearity or continuity. We will have to confront this.
- But first we can define what we want.

Integral kernel systems

Definition

An ***integral kernel system*** is

$$\Sigma = (\mathsf{U},\mathsf{Y},\mathbb{T},\mathscr{U},\mathscr{Y},\mathsf{K}),$$

where

(i) U and Y are finite-dimensional $\mathbb{R}$-vector subspaces,

(ii) $\mathscr{U}$ is a subspace of $\mathsf{C}^0(\mathbb{T};\mathsf{U})$ or $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{T};\mathsf{U})$,

(iii) $\mathscr{Y}$ is a subspace of $\mathsf{C}^0(\mathbb{T};\mathsf{Y})$ or $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{T};\mathsf{Y})$, and

(iv) K is an integral kernel that is compatible with $\mathscr{U}$ and is such that g_{K} is continuous linear mapping into $\mathscr{Y}$.

- We need properties on K, $\mathscr{U}$, and $\mathscr{Y}$ to ensure continuity.
- There is no perfectly general way to do this, so we give a few special cases where this works.

Integral kernel systems

Theorem

Let U *and* Y *be finite-dimensional* $\mathbb{R}$*-vector spaces, let* $\mathbb{T} \subseteq \mathbb{R}$ *be a continuous time-domain. Let* $p \in [1, \infty]$. *Then* $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{K})$ *is an integral kernel system (i.e.,* g_{K} *is continuous) if*

- *(i)* *(a)* $\mathscr{U} \subseteq \mathsf{L}^1(\mathbb{T}; \mathsf{U})$,
  - *(b)* $\mathscr{Y} \subseteq \mathsf{L}^\infty(\mathbb{T}; \mathsf{Y})$, *and*
  - *(c)* *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \mathsf{L}^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\mathsf{L}^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
- *(ii)* *(a)* $\mathscr{U} \subseteq \mathsf{L}^\infty(\mathbb{T}; \mathsf{U})$,
  - *(b)* $\mathscr{Y} \subseteq \mathsf{L}^1(\mathbb{T}; \mathsf{Y})$, *and*
  - *(c)* *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \mathsf{L}^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\mathsf{L}^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
- *(iii)* *(a)* $\mathscr{U} \subseteq \mathsf{L}^p(\mathbb{T}; \mathsf{U})$,
  - *(b)* $\mathscr{Y} \subseteq \mathsf{L}^p(\mathbb{T}; \mathsf{Y})$,
  - *(c)* *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \mathsf{L}^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\mathsf{L}^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and*
  - *(d)* *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \mathsf{L}^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in*

Integral kernel systems

- Causality for an integral kernel system means that, in the expression

$$g_{\mathsf{K}}(\mu)(t) = \int_{\mathbb{T}} \mathsf{K}(t, \tau)\mu(\tau)\mathrm{d}\tau$$

  should only depend on $\mu(\tau)$ for $\tau \leq t$.
- A moment's thought then suggests the following definition and corresponding theorem.

Definition

An integral kernel K is ***causal*** if $\mathsf{K}(t, \tau) = 0$ for $\tau > t$.

Theorem

If Σ *is an integral kernel system with a causal integral kernel* K, *then* g_{K} *is strongly causal.*

- If we use a causal kernel, then we can allow for more general inputs and outputs than the L^p-spaces in the theorem above.

Integral kernel systems

Theorem

Let U *and* Y *be finite-dimensional* $\mathbb{R}$*-vector spaces, let* $\mathbb{T} \subseteq \mathbb{R}$ *be a continuous time-domain. Let* $p \in [1, \infty]$*. Then* $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{K})$ *is an integral kernel system (i.e.,* g_{K} *is continuous) if*

(i) (a) $\mathscr{U} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ *and there is* t_0 *such that* $\inf \operatorname{supp}(\mu) \geq t_0$ *for* $\mu \in \mathscr{U}$,

(b) $\mathscr{Y} \subseteq \mathsf{L}^\infty_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$, *and*

(c) K *is causal and, for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_{\mathbb{K},1}$ *is in* $\mathsf{L}^\infty_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$ *for every compact* $\mathbb{K} \subseteq \mathbb{T}$,

(ii) (a) $\mathscr{U} \subseteq \mathsf{L}^\infty_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ *and there is* t_0 *such that* $\inf \operatorname{supp}(\mu) \geq t_0$ *for* $\mu \in \mathscr{U}$,

(b) $\mathscr{Y} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$, *and*

(c) K *is causal and, for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \mathsf{L}^\infty_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_{\mathbb{K},1}$ *is in* $\mathsf{L}^1_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$ *for every compact* $\mathbb{K} \subseteq \mathbb{T}$,

(iii) (a) $\mathscr{U} \subseteq \mathsf{L}^p_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ *and there is* t_0 *such that* $\inf \operatorname{supp}(\mu) \geq t_0$ *for* $\mu \in \mathscr{U}$,

(b) $\mathscr{Y} \subseteq \mathsf{L}^p_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$,

(c) *(i)(c) and (ii)(c) hold.*

Reading for Lecture 27

1. Section V-6.7.2.

Lecture 28
Continuous-time convolution systems

- Convolution systems arise upon the imposition of stationarity onto integral kernel systems.
- With stationarity, it makes sense to restrict oneself to the time-domain $\mathbb{T} = \mathbb{R}$.

Proposition

Let U *and* Y *be finite-dimensional* $\mathbb{R}$*-vector spaces and let* $\mathsf{K}\colon \mathbb{R} \times \mathbb{R} \to \mathsf{L}(\mathsf{U};\mathsf{Y})$ *be an integral kernel compatible with a set* $\mathscr{U}$ *of input signals. Suppose that* $\mathscr{U}$ *is translation invariant, i.e., that* $\tau_a^* \mu \in \mathscr{U}$ *for every* $a \in \mathbb{R}$ *and* $\mu \in \mathscr{U}$*. Denote*

$$\Sigma_{\mathsf{K}} = (\mathsf{U}, \mathsf{Y}, \mathscr{U}, \mathsf{Y}^{\mathbb{R}}, \mathbb{R}, g_{\mathsf{K}}).$$

Continuous-time convolution systems

Proposition (cont'd)

Then:

(i) if

(a) $\mathsf{K} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}^2; \mathsf{L}(\mathsf{U};\mathsf{Y}))$*,*

(b) $\mathscr{U}$ *has the property that, if* $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ *satisfies*

$$\int_{\mathbb{R}} f(t)\mu(t)\,\mathrm{d}t = 0, \qquad \mu \in \mathscr{U},$$

then $f = 0$*, and*

(c) Σ_{K} *is stationary,*

then there exists $\mathsf{k} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}; \mathsf{L}(\mathsf{U};\mathsf{Y}))$ *such that* $\mathsf{K}(t,\tau) = \mathsf{k}(t-\tau)$ *for almost every* $(t,\tau) \in \mathbb{R}^2$*;*

(ii) if there exists $\mathsf{k} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}; \mathsf{L}(\mathsf{U};\mathsf{Y}))$ *such that* $\mathsf{K}(t,\tau) = \mathsf{k}(t-\tau)$ *for almost every* $(t,\tau) \in \mathbb{R}^2$*, then* Σ_{K} *is strongly stationary.*

Continuous-time convolution systems

- Essentially, then, we see that stationary integral kernel systems have their input/output map defined by

$$g_{\mathsf{K}}(\mu)(t) = \int_{\mathbb{R}} \mathsf{k}(t-\tau)\mu(\tau)\,\mathrm{d}\tau = \mathsf{k} * \mu(t).$$

Definition

A ***continuous-time convolution system*** is

$$\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{R}, \mathscr{U}, \mathscr{Y}, \mathsf{k}),$$

where

(i) U and Y are finite-dimensional $\mathbb{R}$-vector spaces,

(ii) $\mathscr{U}$ is a subspace of $\mathsf{C}^0(\mathbb{T};\mathsf{U})$ or $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{T};\mathsf{U})$,

(iii) $\mathscr{Y}$ is a subspace of $\mathsf{C}^0(\mathbb{T};\mathsf{Y})$ or $\mathsf{L}^p_{\mathrm{loc}}(\mathbb{T};\mathsf{Y})$, and

(iv) $\mathsf{k}\colon \mathbb{R} \to \mathsf{L}(\mathsf{U};\mathsf{Y})$

Continuous-time convolution systems

Definition (cont'd)

are such that, if we take $\mathsf{K}(t,\tau) = \mathsf{k}(t-\tau)$, then

$$\Sigma' = (\mathsf{U}, \mathsf{Y}, \mathbb{R}, \mathscr{U}, \mathscr{Y}, \mathsf{K})$$

is an integral kernel system.

- We call k a ***convolution kernel***.
- The notion of causality for integral kernel systems transfers easily to continuous-time convolution systems.

Definition

A continuous-time convolution kernel

$$\mathsf{k}\colon \mathbb{R} \to \mathsf{L}(\mathsf{U};\mathsf{Y})$$

is ***causal*** if $\mathsf{k}(t) = 0$ for $t \in \mathbb{R}_{<0}$.

Continuous-time convolution systems

- As with integral kernel systems, one must have conditions on $\mathscr{U}$, $\mathscr{Y}$, and k to ensure continuity of the input/output map.
- One can convert the conditions we have for integral kernel systems, but we have already carefully considered the matter of continuity of convolution on Slide 92. In the causal case, we restrict ourselves to signals that are zero for native time. Thus we have continuity of convolution kernel systems in the following cases:
  1. $\mathscr{U} \subseteq \mathsf{L}^1(\mathbb{R};\mathsf{U})$, $\mathscr{Y} \subseteq \mathsf{L}^1(\mathbb{R};\mathsf{Y})$, and $\mathsf{k} \in \mathsf{L}^1(\mathbb{R};\mathsf{L}(\mathsf{U};\mathsf{Y}))$;
  2. $\mathscr{U} \subseteq \mathsf{L}^p(\mathbb{R};\mathsf{U})$, $\mathscr{Y} \subseteq \mathsf{L}^q(\mathbb{R};\mathsf{Y})$, and $\mathsf{k} \in \mathsf{L}^r(\mathbb{R};\mathsf{L}(\mathsf{U};\mathsf{Y}))$, where $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}$;
  3. $\mathscr{U} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0};\mathsf{U})$, $\mathscr{Y} \subseteq \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0};\mathsf{Y})$, and $\mathsf{k} \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R}_{\geq 0};\mathsf{L}(\mathsf{U};\mathsf{Y}))$;
  4. $\mathscr{U} \subseteq \mathsf{L}^p_{\mathrm{loc}}(\mathbb{R}_{\geq 0};\mathsf{U})$, $\mathscr{Y} \subseteq \mathsf{L}^q_{\mathrm{loc}}(\mathbb{R}_{\geq 0};\mathsf{Y})$, and $\mathsf{k} \in \mathsf{L}^r_{\mathrm{loc}}(\mathbb{R}_{\geq 0};\mathsf{L}(\mathsf{U};\mathsf{Y}))$, where $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}$.

Linear continuous-time input/output systems from linear continuous-time state space systems

- On Slide 129 we initiated a programme to produce continuous-time input/output systems from continuous-time state space systems. We saw that there were substantial technical considerations.
- Here we do the same for linear systems, where things are quite a lot easier.
- We shall consider the input/output systems arising from that part of the output equations we had labelled as *term 2*, i.e., the parts coming from the proper impulse transmission map (in the time-varying case) and the proper impulse response (in the constant coefficient case).
- As we saw in the not necessarily linear case, one must take care of the fact that the input/output map is continuous, and this requires considering various cases of input, output, and system assumptions.

Linear continuous-time input/output systems from linear continuous-time state space systems

Theorem

Let $\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$ *be a linear continuous-time state space system and let* $p \in [1, \infty]$*. Let* $t_0 \in \mathbb{T}$ *and let*

$$\mathscr{U} \subseteq \{\mu \in \mathsf{L}^p_{\mathsf{loc}}(\mathbb{T}; \mathsf{U}) \mid \mu(t) = 0,\ t < t_0\},$$
$$\mathscr{Y} = \{\eta \in \mathsf{L}^p_{\mathsf{loc}}(\mathbb{T}; \mathsf{Y}) \mid \eta(t) = 0,\ t < t_0\}.$$

Then

$$\Sigma_{\mathsf{i/o}}(t_0) = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{pitm}_\Sigma)$$

is a causal integral kernel system in the following cases:

(i) *(a)* $\mathsf{B} \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{X}))$ *and* $\mathsf{C} \in \mathsf{L}^\infty_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{Y}))$ *and*
(b) $p = \infty$*;*

(ii) *(a)* $\mathsf{B} \in \mathsf{L}^\infty_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{X}))$ *and* $\mathsf{C} \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{Y}))$ *and*
(b) $p = 1$*;*

Linear continuous-time input/output systems from linear continuous-time state space systems

Theorem (cont'd)

(iii) *(a)* $\mathsf{B} \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{X}))$ *and* $\mathsf{C} \in \mathsf{L}^\infty_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{Y}))$*,*
(b) $\mathsf{B} \in \mathsf{L}^\infty_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{X}))$ *and* $\mathsf{C} \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{Y}))$*, and*
(c) $p \in [1, \infty]$*.*

- Of course, the conditions come directly from the conditions for continuity of the input/output map from integral kernel systems with a causal integral kernel, which pitm_Σ is.
- To state the corresponding result in the constant coefficient case follows along similar lines, adapting the conditions for a causal continuous-time convolution system to be continuous.

Linear continuous-time input/output systems from linear continuous-time state space systems

Theorem

Let $\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{R}, \mathscr{U}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$ *be a linear continuous-time state space system with constant coefficients and let* $p, q, r \in [1, \infty]$ *satisfy one of the following two criterion: (1)* $p = q = r = 1$*; (2)* $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}$*. Let*

$$\mathscr{U} \subseteq \{\mu \in \mathsf{L}^p_{\mathrm{loc}}(\mathbb{R}; \mathsf{U}) \mid \mu(t) = 0,\ t < 0\},$$
$$\mathscr{Y} = \{\eta \in \mathsf{L}^q_{\mathrm{loc}}(\mathbb{R}; \mathsf{Y}) \mid \eta(t) = 0,\ t < 0\}.$$

Then

$$\Sigma_{\mathrm{i/o}} = (\mathsf{U}, \mathsf{Y}, \mathbb{R}, \mathscr{U}, \mathscr{Y}, \mathrm{pir}_\Sigma)$$

is a causal continuous-time convolution system.

- *Punchline:* Integral kernel systems and continuous-time convolutions systems arise in a natural way from linear continuous-time state space systems, and this explains, in part, their importance.

Reading for Lecture 28

1. Section V-6.7.4.
2. Section V-6.7.6.

Lecture 29

Aside: The generality of integral kernel systems and continuous-time convolution systems

- We have presented integral kernel systems and their stationary sisters, continuous-time convolution systems, as examples of linear continuous-time input/output systems.
- *Question:* Are there linear continuous-time input systems that are not integral kernel systems or continuous-time convolution systems.
- *Answer:* Yes. . . but *a lot* of linear continuous-time input systems *are* integral kernel systems or continuous-time convolution systems.
- To be precise about this requires two things:
  1. relaxing what one means by an integral kernel system or continuous-time convolution system;
  2. taking advantage of a powerful theorem in the theory of distributions called the "Schwartz Kernel Theorem."
- For the Schwartz Kernel Theorem, we need two things:
  1. one can define test functions $\mathscr{D}(\mathbb{R}^n;\mathbb{F})$ and corresponding distributions $\mathscr{D}'(\mathbb{R}^n;\mathbb{F})$ in multiple variables, rather like we did in one variable;
  2. if $\phi,\psi\in\mathscr{D}(\mathbb{R};\mathbb{F})$, define $\phi\otimes\psi\in\mathscr{D}(\mathbb{R}^2;\mathbb{F})$ by $\phi\otimes\psi(x,y)=\phi(x)\psi(y)$.

Aside: The generality of integral kernel systems and continuous-time convolution systems

Theorem

If $\Phi\colon\mathscr{D}(\mathbb{R};\mathsf{U})\to\mathscr{D}'(\mathbb{R};\mathsf{Y})$ is a continuous linear map, then there exists a distribution $\mathsf{K}\in\mathscr{D}'(\mathbb{R}^2;\mathsf{L}(\mathsf{U};\mathsf{Y}))$ such that

$$\langle\mathsf{K};\phi\otimes\psi\rangle=\langle\Phi(\phi);\psi\rangle,\qquad\phi\in\mathscr{D}(\mathbb{R};\mathsf{U}),\ \psi\in\mathscr{D}(\mathbb{R};\mathsf{Y}).$$

- The way to read the theorem is that Φ is defined by "integrating" with respect to τ in $(t,\tau)\in\mathbb{R}^2$, but integration means using the distribution K.
- Let us now see how to use this rather abstract-seeming theorem in system theory.
  1. Let $\mathscr{U}\subseteq\mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathsf{U})$ be such that, if $\mu\in\mathscr{U}$, then there is a sequence $(\phi_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathscr{D}(\mathbb{R};\mathsf{U})$ converging to μ in the topology of $\mathscr{U}$. For example, one can take $\mathscr{U}\subseteq\mathsf{L}^p(\mathbb{R};\mathsf{U})$, $p\in[1,\infty)$, or $\mathscr{U}\subseteq\mathsf{C}^0_0(\mathbb{R};\mathsf{U})$. These are easily believed to work because they were *defined* to be the completion of $\mathsf{C}^0_{\mathrm{cpt}}(\mathbb{R};\mathsf{U})$. However, we cannot use $\mathscr{U}\subseteq\mathsf{L}^\infty(\mathbb{R};\mathsf{U})$ or $\mathscr{U}\subseteq\mathsf{C}^0_{\mathrm{bdd}}(\mathbb{R};\mathsf{U})$.

Aside: The generality of integral kernel systems and continuous-time convolution systems

2. Let $\mathscr{Y}$ be such that there is a continuous inclusion

$$\mathscr{Y} \ni \eta \mapsto \theta_\eta \in \mathscr{D}'(\mathbb{R}; \mathsf{Y}).$$

   This holds for *all* of the continuous-time signal spaces we have used, either in MATH/MTHE 334 or in this course.
3. We thus have the following sequence of continuous linear mappings

$$\mathscr{D}(\mathbb{R}; \mathsf{U}) \longrightarrow \mathscr{U} \xrightarrow{\Phi} \mathscr{Y} \longrightarrow \mathscr{D}'(\mathbb{R}; \mathsf{Y})$$

4. Now suppose that we have a continuous linear map $\Phi\colon \mathscr{U} \to \mathscr{Y}$. We then have a mapping $\hat{\Phi}\colon \mathscr{D}(\mathbb{R}; \mathsf{U}) \to \mathscr{Y}$ by restriction of Φ to $\mathscr{D}(\mathbb{R}; \mathsf{U})$.
5. Then, by the Schwartz Kernel Theorem, there is $\mathsf{K} \in \mathscr{D}(\mathbb{R}^2; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$ such that

$$\langle \mathsf{K}; \phi \otimes \psi \rangle = \langle \hat{\Phi}(\phi); \psi \rangle, \phi \in \mathscr{D}(\mathbb{R}; \mathsf{U}), \ \psi \in \mathscr{D}(\mathbb{R}; \mathsf{Y}).$$

Aside: The generality of integral kernel systems and continuous-time convolution systems

6. By continuity and since, for every $\mu \in \mathscr{U}$, there is a sequence $(\phi_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathscr{D}(\mathbb{R}; \mathsf{U})$ converging to μ, we have

$$\langle \theta_{\Phi(\mu)}; \psi \rangle = \lim_{j \to \infty} \langle \mathsf{K}; \phi_j \otimes \psi \rangle, \qquad \psi \in \mathscr{D}(\mathbb{R}; \mathsf{Y}).$$

- *Punchline:* A large number of linear continuous-time input/output systems are integral kernel systems, provided that you allow distributions as integral kernels.
- *Corollary to punchline:* A large number of stationary linear continuous-time input/output systems are continuous-time convolution systems, provided that you allow distributions as convolution kernels.

Linear discrete-time state space systems

- We now carry out the programme for linear discrete-time state space systems that was carried out previously for linear continuous-time state space systems.

Definition

A ***linear discrete-time state space system*** is

$$\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathbb{T}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}),$$

where

(i) X is a finite-dimensional $\mathbb{R}$-vector space (***state space***),

(ii) U is a finite-dimensional $\mathbb{R}$-vector space (***input set***),

(iii) Y is a finite-dimensional $\mathbb{R}$-vector space (***output set***),

(iv) $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ is a discrete time-domain (***time-domain***),

(v) $\mathscr{U} \subseteq \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ (***inputs***),

(vi) $\mathsf{A} \in \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{X}))$,

(vii) $\mathsf{B} \in \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{X}))$,

(viii) $\mathsf{C} \in \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{X}; \mathsf{Y}))$, and

(ix) $\mathsf{D} \in \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$.

Linear discrete-time state space systems

- A linear discrete-time state space system is an instance of a discrete-time state space system by defining the dynamics and the output map by

$$f(t, x, u) = \mathsf{A}(t)x + \mathsf{B}(t)u, \quad h(t, x, u) = \mathsf{C}(t)x + \mathsf{D}(t)u.$$

  respectively.
- Note that, for linear systems, we do not fuss with partially defined inputs, state, and outputs, although one *can* do this.
- We see, then, that $(\xi, \mu) \in \mathrm{Ctraj}(\Sigma)$ if and only if

$$\xi(t + \Delta) = \mathsf{A}(t) \cdot \xi(t) + \mathsf{B}(t) \cdot \mu(t).$$

- Unlike in the nonlinear case, or even the linear continuous-time case, there is no existence and uniqueness theorem.
  1. Trajectories exist.
  2. They are not, in general unique, owing to the system possibly not being invertible. In this case, invertibility amounts to the invertibility of $\mathsf{A}(t)$, $t \in \mathbb{T}$.

Linear discrete-time state space systems

- Let us carefully develop the flow for controlled trajectories for a linear discrete-time state space system, as we did not do this when we talked about difference equations.
- First consider the homogeneous equation

$$\xi(t+\Delta) = \mathsf{A}(t)\cdot\xi(t), \qquad \xi(t_0) = x_0.$$

Just by recursion:

$$\begin{aligned}
\xi(t_0) &= x_0 \\
\xi(t_0+\Delta) &= \mathsf{A}(t_0)(x_0) \\
\xi(t_0+2\Delta) &= \mathsf{A}(t_0+\Delta)\circ\mathsf{A}(t_0)(x_0) \\
&\vdots \\
\xi(t_0+j\Delta) &= \mathsf{A}(t_0+(j-1)\Delta)\circ\cdots\circ\mathsf{A}(t_0)(x_0) \\
&\vdots
\end{aligned}$$

Linear discrete-time state space systems

- We thus define $\Phi^{\mathsf{d}}_{\mathsf{A},t_0} = \mathsf{A}(t_0+(j-1)\Delta)\circ\cdots\circ\mathsf{A}(t_0)$, this being the state transition map in the discrete-time case.
- Now we have the variation of constants formula in the discrete-time case:

$$\Phi^{\Sigma}(t,t_0,x_0,\mu) = \Phi^{\mathsf{d}}_{\mathsf{A},t_0}(t)(x_0) + \sum_{j=0}^{(t-t_0-\Delta)/\Delta} \Phi^{\mathsf{d}}_{\mathsf{A},t_0+(j+1)\Delta}(t)\circ\mathsf{B}(t_0+j\Delta)(\mu(t_0+j\Delta)).$$

- The corresponding output is then

$$\begin{aligned}
\eta(t) = {} & \underbrace{\mathsf{C}(t)\circ\Phi^{\mathsf{d}}_{\mathsf{A}}(t,t_0)(x_0)}_{\textit{term 1}} \\
& + \underbrace{\sum_{j=0}^{(t-t_0-\Delta)/\Delta} \mathsf{C}(t)\circ\Phi^{\mathsf{d}}_{\mathsf{A},t_0+(j+1)\Delta}(t)\circ\mathsf{B}(t_0+j\Delta)(\mu(t_0+j\Delta))}_{\textit{term 2}} + \underbrace{\mathsf{D}(t)\cdot\mu(t)}_{\textit{term 3}}.
\end{aligned}$$

Linear discrete-time state space systems

- We note that the output is comprised of three bits, each interesting in its own right:
  1. *term 1*: Here is a component of the output determined by the initial condition. Indeed, the other parts of the solution are independent of initial condition.
  2. *term 2*: This is some weird sum. We shall encounter this sort of thing when we consider linear discrete-time input/output systems below.
  3. *term 3*: This term consists of the input at time t directly influencing the output at time t in a memoryless fashion. This is sometimes called a "feedforward" term.

Linear discrete-time state space systems

- Next we consider systems with constant coefficients.

Definition

A ***linear discrete-time state space system with constant coefficients*** is

$$\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathbb{T}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}),$$

where

(i) X is a finite-dimensional $\mathbb{R}$-vector space (***state space***),

(ii) U is a finite-dimensional $\mathbb{R}$-vector space (***input set***),

(iii) Y is a finite-dimensional $\mathbb{R}$-vector space (***output set***),

(iv) $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ is a discrete time-domain (***time-domain***),

(v) $\mathscr{U} \subseteq \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ (***inputs***),

(vi) $\mathsf{A} \in \mathsf{L}(\mathsf{X}; \mathsf{X})$,

(vii) $\mathsf{B} \in \mathsf{L}(\mathsf{U}; \mathsf{X})$,

(viii) $\mathsf{C} \in \mathsf{L}(\mathsf{X}; \mathsf{Y})$, and

(ix) $\mathsf{D} \in \mathsf{L}(\mathsf{U}; \mathsf{Y})$.

Linear discrete-time state space systems

- In short, a "linear discrete time state space system with constant coefficients" is a linear discrete-time state space system where A, B, C, and D are independent of time.
- As such, a linear discrete-time state space system with constant coefficients is also an autonomous discrete-time state space system with

$$f(x, u) = \mathsf{A}(x) + \mathsf{B}(u), \quad h(x, u) = \mathsf{C}(x) + \mathsf{D}(u).$$

- Thus a controlled trajectory $(\xi, \mu) \in \mathsf{Ctraj}(\Sigma)$ with corresponding controlled output $(\eta, \mu) \in \mathsf{Cout}(\Sigma)$ jointly satisfy

$$\begin{aligned} \xi(t + \Delta) &= \mathsf{A} \circ \xi(t) + \mathsf{B} \circ \mu(t), \\ \eta(t) &= \mathsf{C} \circ \xi(t) + \mathsf{D} \circ \mu(t). \end{aligned}$$

- Note that, when A is independent of time,

$$\Phi^{\mathsf{d}}_{\mathsf{A}}(t_0 + j\Delta, t_0) = \mathsf{P}_{\mathsf{A}}(j),$$

where $\mathsf{P}_{\mathsf{A}} \colon \mathbb{Z}_{\geq 0} \to \mathsf{L}(\mathsf{X}; \mathsf{X})$ is defined by $\mathsf{P}_{\mathsf{A}}(j) = \mathsf{A}^j$.

Linear discrete-time state space systems

- Thus the formula for the flow and the output equations is as in the non-constant coefficient case, but making this substitution for the state transition map.
- Thus we have

$$\Phi^{\Sigma}(t, t_0, x_0) = \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-t_0}{\Delta}\right)(x_0) + \sum_{j=0}^{(t-t_0-\Delta)/\Delta} \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-t_0-(j+1)\Delta}{\Delta}\right)(\mathsf{B}(\mu(t_0 + j\Delta)))$$

and

$$\begin{aligned} \eta(t) = {} & \mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-t_0}{\Delta}\right)(x_0) \\ & + \sum_{j=0}^{(t-t_0-\Delta)/\Delta} \mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-t_0-(j+1)\Delta}{\Delta}\right)(\mathsf{B}(\mu(t_0 + j\Delta))) + \mathsf{D} \circ \mu(t). \end{aligned}$$

Linear discrete-time state space systems

- Because linear continuous-time state space systems with constant coefficients are stationary whenever the input set $\mathscr{U}$ is translation-invariant, one commonly makes two assumptions:
  1. $\mathbb{T} = \mathbb{Z}(\Delta)$;
  2. one considers the initial time to be $t_0 = 0$.
- As a consequence of this, one typically works with inputs defined on $\mathbb{Z}_{\geq 0}(\Delta)$ and with controlled trajectories and outputs determined by

$$\Phi^{\Sigma}(t, 0, x_0) = \mathsf{P}_{\mathsf{A}}\left(\tfrac{t}{\Delta}\right)(x_0) + \sum_{j=0}^{(t-\Delta)/\Delta} \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-(j+1)\Delta}{\Delta}\right)(\mathsf{B}(\mu(j\Delta)))$$

and

$$\eta(t) = \underbrace{\mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t}{\Delta}\right)(x_0)}_{\textit{term 1}} + \underbrace{\sum_{j=0}^{(t-\Delta)/\Delta} \mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-(j+1)\Delta}{\Delta}\right)(\mathsf{B}(\mu(j\Delta)))}_{\textit{term 2}} + \underbrace{\mathsf{D} \circ \mu(t)}_{\textit{term 3}}.$$

- The three terms have a similar interpretation as in the time-varying case, except that *term 2* is a convolution. We shall care about this.

Reading for Lecture 29

1. Sections V-6.7.3 and V-6.7.5 (stay tuned. . .).
2. Sections V-6.8.1 and V-6.8.2.

Lecture 30

Linear discrete-time state space systems (cont'd)

- For linear discrete-time state space systems, we had the following formulae for outputs.
  1. Time-varying case, input μ, initial condition x_0 at initial time t_0:
  2. The corresponding output is then

$$\eta(t) = \underbrace{\mathsf{C}(t) \circ \Phi^{\mathsf{d}}_{\mathsf{A}}(t, t_0)(x_0)}_{\textit{term 1}} + \underbrace{\sum_{j=0}^{(t-t_0-\Delta)/\Delta} \mathsf{C}(t) \circ \Phi^{\mathsf{d}}_{\mathsf{A}, t_0+(j+1)\Delta}(t) \circ \mathsf{B}(t_0 + j\Delta)(\mu(t_0 + j\Delta))}_{\textit{term 2}} + \underbrace{\mathsf{D}(t) \cdot \mu(t)}_{\textit{term 3}}.$$

  3. Constant coefficient case, input μ, initial condition x_0 at initial time 0:

$$\eta(t) = \underbrace{\mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t}{\Delta}\right)(x_0)}_{\textit{term 1}} + \underbrace{\sum_{j=0}^{(t-\Delta)/\Delta} \mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-(j+1)\Delta}{\Delta}\right)(\mathsf{B}(\mu(j\Delta)))}_{\textit{term 2}} + \underbrace{\mathsf{D} \circ \mu(t)}_{\textit{term 3}}.$$

- Our objective now is to explore more fully the components labelled in the preceding formulae as "*term 2*."

Linear discrete-time state space systems

- Let $\mathsf{P}\colon \mathbb{Z}(\Delta) \to \mathbb{R}$ be the pulse at $t = 0$.

Definition

Let σ be a linear discrete-time state space system.

(i) The ***proper impulse transmission map*** for Σ at t_0 is the function $\mathsf{pitm}_{\Sigma,t_0}\colon \mathbb{T} \to \mathsf{L}(\mathsf{U};\mathsf{Y})$ defined by

$$\mathsf{pitm}_{\Sigma,t_0}(t) = \mathsf{1}_{\geq 0}(t - (t_0 + \Delta))\mathsf{C}(t) \circ \Phi^{\mathsf{d}}_{\mathsf{A},t_0+\Delta}(t) \circ \mathsf{B}(t_0).$$

(ii) The ***impulse transmission map*** for Σ at $t_0 \in \mathbb{T}$ is the function $\mathsf{itm}_{\Sigma,t_0}\colon \mathbb{T} \to \mathsf{L}(\mathsf{U};\mathsf{Y})$ defined by

$$\mathsf{itm}_{\Sigma,t_0}(t) = \mathsf{pitm}_{\Sigma,t_0}(t) + \tau^*_{t_0}\mathsf{P}(t)\mathsf{D}(t).$$

- Obviously, the output for an input μ and initial condition x_0 at t_0 is

$$\eta(t) = \mathsf{C}(t) \circ \Phi^{\mathsf{d}}_{\mathsf{A}}(t, t_0)(x_0) + \sum_{j=0}^{(t-t_0-\Delta)/\Delta} \mathsf{pitm}_{\Sigma,t_0+j\Delta}(t)\mu(t_0 + j\Delta) + \mathsf{D}(t) \cdot \mu(t). \quad (9)$$

Linear discrete-time state space systems

- Let us make a few comments about the preceding.
  1. When $\mathsf{D}(t) = 0$ for all t and when $x_0 = 0$,

$$\mathsf{pitm}_{\Sigma,t_0}(t) = \sum_{j=0}^{(t-t_0-\Delta)/\Delta} \mathsf{pitm}_{\Sigma,t_0+j\Delta}(t)\tau_{t_0}^{*}\mathsf{P}(t_0 + j\Delta)$$

     gives the output for an input that is a pulse at t_0.
  2. The formula also gives context to the terminology "impulse transmission map." We see that $\mathsf{pitm}_{\Sigma,t_0+j\Delta}(t)$ in the summand serves to "transmit" the effect of the input at time $t_0 + j\Delta$ to the output at time t.
- Now we look at the impulse transmission map in the constant coefficient case.
- Unlike in the continuous-time case, we do not have to mess about with distributions.

Linear discrete-time state space systems

Definition

Let Σ be a linear discrete-time state space system with constant coefficients.

(i) The ***proper impulse response*** for Σ is the function

$$\begin{aligned}\mathsf{pir}_{\Sigma} \colon \mathbb{Z}(\Delta) &\to \mathsf{L}(\mathsf{U};\mathsf{Y})\\ t &\mapsto 1_{\geq 0}(t-\Delta)\mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t-\Delta}{\Delta}\right) \circ \mathsf{B}.\end{aligned}$$

(ii) The ***impulse response*** for Σ is the function

$$\begin{aligned}\mathsf{ir}_{\Sigma} \colon \mathbb{Z}(\Delta) &\to \mathsf{L}(\mathsf{U};\mathsf{Y})\\ t &\mapsto \mathsf{pir}_{\Sigma}(t) + \mathsf{P}(t)\mathsf{D}.\end{aligned}$$

Linear discrete-time state space systems

- Note that a linear discrete-time state space system with constant coefficients is, in particular, a linear discrete-time state space system. As such, in the constant coefficient case, we still have the notion of an impulse transmission map. Indeed, we have

$$\mathsf{pitm}_{\Sigma,\tau}(t) = \mathsf{pir}_{\Sigma}(t-\tau), \qquad t \geq \tau.$$

  in the constant coefficient case.
- Note that the output associated to the input μ with initial condition x_0 at $t = 0$ is

$$\eta(t) = \mathsf{C} \circ \mathsf{P}_{\mathsf{A}}\left(\tfrac{t}{\Delta}\right)(x_0) + \sum_{j=0}^{(t-\Delta)/\Delta} \mathsf{pir}_{\Sigma}\left(\tfrac{t-j\Delta}{\Delta}\right)(\mu(j\Delta)) + \mathsf{D} \circ \mu(t). \qquad (10)$$

- The middle term is an old friend, namely a convolution!
- Summary:
  1. Both the impulse transmission map (in the time-varying case) and the impulse response (in the constant coefficient case) are the output for a pulse input.
  2. The formulae (9) and (10) illustrate that this response to an impulse forms an integral (less effective pun) part of the output for a general input.

Linear discrete-time input/output systems

- To use linearity, we need to assume that $\mathscr{U}$ and $\mathscr{Y}$ are such that linearity from $\mathscr{U}$ to $\mathscr{Y}$ makes sense.
- Recall that, if $\mathscr{U} \subseteq \ell_{\mathrm{loc}}((\mathbb{T}); U)$, then we had denoted, for $\mathbb{S} \subseteq \mathbb{T}$,

$$\mathscr{U}(\mathbb{S}) = \{\mu \in \mathscr{U} \mid \operatorname{dom}(\mu) = \mathbb{S}\}.$$

Definition

An ***linear discrete-time input/output system*** is $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, g)$, where

(i) U is a finite dimensional $\mathbb{R}$-vector space (***input set***),

(ii) Y is a finite dimensional $\mathbb{R}$-vector space (***output set***),

(iii) $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ is a discrete time-domain,

(iv) $\mathscr{U}$ is such that, for every $\mathbb{S} \subseteq \mathbb{T}$, $\mathscr{U}(\mathbb{S})$ is a subspace of $\mathsf{U}^{\mathbb{S}}$,

(v) $\mathscr{Y}$ is such that, for every $\mathbb{S} \subseteq \mathbb{T}$, $\mathscr{Y}(\mathbb{S})$ is a subspace of $\mathsf{Y}^{\mathbb{S}}$, and

Linear discrete-time input/output systems

Definition (cont'd)

(vi) (a) for $\mathbb{S} \subseteq \mathbb{T}$, if $g_{\mathbb{S}} = g|\mathscr{U}(\mathbb{S})$, then $g_{\mathbb{S}}(\mu) \in \mathscr{Y}(\mathbb{S})$,

(b) if $\mathbb{S}' \subseteq \mathbb{S} \subseteq \mathbb{T}$, then

$$g_{\mathbb{S}'}(\mu|\mathbb{S}') = g_{\mathbb{S}}(\mu)|\mathbb{S}',$$

and

(c) for $\mathbb{S} \subseteq \mathbb{T}$, $g_{\mathbb{S}} \colon \mathscr{U}(\mathbb{S}) \to \mathscr{Y}(\mathbb{S})$ is a continuous linear mapping.

- In brief, a linear discrete-time input/output system is a discrete-time input/output system that is. . . linear.

Summation kernel systems

- The basic idea:

$$g_{\mathsf{K}}(\mu)(t) = \sum_{\tau \in \mathbb{T}} \mathsf{K}(t, \tau)\mu(\tau).$$

We call K the "summation kernel."
- This should remind you of the proper impulse transmission map for linear discrete-time state space systems.

Definition

Let U, Y be finite-dimensional $\mathbb{R}$-vector spaces, and let $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ be a discrete time-domain.

(i) A ***summation kernel*** is a mapping

$$\mathsf{K} \colon \mathbb{T} \times \mathbb{T} \to \mathsf{L}(\mathsf{U}; \mathsf{Y}).$$

(ii) We denote

$$\begin{aligned} \mathsf{K}_t \colon \mathbb{T} &\to \mathsf{L}(\mathsf{U}; \mathsf{Y}) \\ \tau &\mapsto \mathsf{K}(t, \tau). \end{aligned}$$

Summation kernel systems

Definition (cont'd)

(iii) If $\mathscr{U} \subseteq \mathsf{U}^{\mathbb{T}}$, then K is ***compatible*** with $\mathscr{U}$ if $\tau \mapsto \mathsf{K}_t(\tau)\mu(\tau)$ is in $\ell^1(\mathbb{T}; \mathsf{Y})$ for every $\mu \in \mathscr{U}$.

(iv) if K is compatible with $\mathscr{U} \subseteq \mathsf{U}^{\mathbb{T}}$, then the ***summation operator*** associated with K is

$$g_{\mathsf{K}} : \mathscr{U} \to \mathsf{Y}^{\mathbb{T}}$$

defined by

$$g_{\mathsf{K}}(\mu)(t) = \sum_{\tau \in \mathbb{T}} \mathsf{K}(t, \tau)\mu(\tau).$$

- At this point, we cannot quite call this a linear discrete-time input/output system since we do not have linearity or continuity. We will have to confront this.
- But first we can define what we want.

Summation kernel systems

Definition

A ***summation kernel system*** is

$$\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{K}),$$

where

(i) U and Y are finite-dimensional $\mathbb{R}$-vector subspaces,

(ii) $\mathscr{U}$ is a subspace of $\ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$,

(iii) $\mathscr{Y}$ is a subspace of $\ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$, and

(iv) K is a summation kernel that is compatible with $\mathscr{U}$ and is such that g_{K} is continuous linear mapping into $\mathscr{Y}$.

- We need properties on K, $\mathscr{U}$, and $\mathscr{Y}$ to ensure continuity.
- There is no perfectly general way to do this, so we give a few special cases where this works.

Summation kernel systems

Theorem

Let U *and* Y *be finite-dimensional* $\mathbb{R}$*-vector spaces, let* $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$ *be a discrete time-domain. Let* $p \in [1, \infty]$*. Then* $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{K})$ *is a summation kernel system (i.e.,* g_{K} *is continuous) if*

(i) *(a)* $\mathscr{U} \subseteq \ell^1(\mathbb{T}; \mathsf{U})$,
(b) $\mathscr{Y} \subseteq \ell^\infty(\mathbb{T}; \mathsf{Y})$, *and*
(c) *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \ell^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\ell^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,

(ii) *(a)* $\mathscr{U} \subseteq \ell^\infty(\mathbb{T}; \mathsf{U})$,
(b) $\mathscr{Y} \subseteq \ell^1(\mathbb{T}; \mathsf{Y})$, *and*
(c) *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \ell^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\ell^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,

(iii) *(a)* $\mathscr{U} \subseteq \ell^p(\mathbb{T}; \mathsf{U})$,
(b) $\mathscr{Y} \subseteq \ell^p(\mathbb{T}; \mathsf{Y})$,
(c) *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \ell^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\ell^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and*
(d) *for each* $t \in \mathbb{T}$, $\mathsf{K}_t \in \ell^\infty(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, *and* $t \mapsto \|\mathsf{K}_t\|_1$ *is in* $\ell^1(\mathbb{T}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$.

Summation kernel systems

- Causality for a summation kernel system means that, in the expression

$$g_{\mathsf{K}}(\mu)(t) = \sum_{\tau \in \mathbb{T}} \mathsf{K}(t, \tau)\mu(\tau)$$

should only depend on $\mu(\tau)$ for $\tau \leq t$.

- A moment's thought then suggests the following definition and corresponding theorem.

Definition

A summation kernel K is ***causal*** if $\mathsf{K}(t, \tau) = 0$ for $\tau > t$.

Theorem

If Σ *is a summation kernel system with a causal summation kernel* K*, then* g_{K} *is causal.*

- If we use a causal kernel, then we can allow for more general inputs and outputs than the ℓ^p-spaces in the theorem above.

Summation kernel systems

Theorem

Let U *and* Y *be finite-dimensional* $\mathbb{R}$*-vector spaces, let* $\mathbb{T} \subseteq \mathbb{R}$ *be a continuous time-domain. Let* $p \in [1, \infty]$*. Then* $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{K})$ *is an integral kernel system (i.e.,* g_{K} *is continuous) if*

(i) $\mathscr{U} \subseteq \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ *and there is* t_0 *such that* $\inf \operatorname{supp}(\mu) \geq t_0$ *for* $\mu \in \mathscr{U}$*,*

(ii) $\mathscr{Y} \subseteq \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$*, and*

(iii) K *is causal.*

Reading for Lecture 30

1. Section V-6.8.3.
2. Sections V-6.9.1 and V-6.9.2.

Lecture 31
Discrete-time convolution systems

- Convolution systems arise upon the imposition of stationarity onto summation kernel systems.
- With stationarity, it makes sense to restrict oneself to the time-domain $\mathbb{T} = \mathbb{Z}(\Delta)$.

Proposition

Let U *and* Y *be finite-dimensional* $\mathbb{R}$*-vector spaces and let* $\mathsf{K}\colon \mathbb{Z}(\Delta) \times \mathbb{Z}(\Delta) \to \mathsf{L}(\mathsf{U};\mathsf{Y})$ *be an summation kernel compatible with a set* $\mathscr{U}$ *of input signals. Suppose that* $\mathscr{U}$ *is translation invariant, i.e., that* $\tau_a^*\mu \in \mathscr{U}$ *for every* $a \in \mathbb{Z}(\Delta)$ *and* $\mu \in \mathscr{U}$*. Denote*

$$\Sigma_{\mathsf{K}} = (\mathsf{U}, \mathsf{Y}, \mathscr{U}, \mathsf{Y}^{\mathbb{Z}(\Delta)}, \mathbb{Z}(\Delta), g_{\mathsf{K}}).$$

Discrete-time convolution systems

Proposition (cont'd)

Then:

(i) if

(a) $\mathscr{U}$ *has the property that, if* $f \in \ell_{\mathrm{loc}}(\mathbb{Z}(\Delta);\mathbb{R})$ *satisfies*

$$\sum_{t\in\mathbb{Z}(\Delta)} f(t)\mu(t)\,\mathrm{d}t = 0, \qquad \mu \in \mathscr{U},$$

then $f = 0$*, and*

(b) Σ_{K} *is stationary,*

then there exists $\mathsf{k} \in \ell_{\mathrm{loc}}(\mathbb{Z}(\Delta);\mathsf{L}(\mathsf{U};\mathsf{Y}))$ *such that* $\mathsf{K}(t,\tau) = \mathsf{k}(t-\tau)$ *for almost every* $(t,\tau) \in \mathbb{Z}(\Delta)^2$*;*

(ii) if there exists $\mathsf{k} \in \ell_{\mathrm{loc}}(\mathbb{Z}(\Delta);\mathsf{L}(\mathsf{U};\mathsf{Y}))$ *such that* $\mathsf{K}(t,\tau) = \mathsf{k}(t-\tau)$ *for almost every* $(t,\tau) \in \mathbb{Z}(\Delta)^2$*, then* Σ_{K} *is strongly stationary.*

Discrete-time convolution systems

- Essentially, then, we see that stationary summation kernel systems have their input/output map defined by

$$g_{\mathsf{K}}(\mu)(t) = \sum_{\tau \in \mathbb{Z}(\Delta)} \mathsf{k}(t-\tau)\mu(\tau) = \mathsf{k} * \mu(t).$$

Definition

A ***discrete-time convolution system*** is

$$\Sigma = (\mathsf{U}, \mathsf{Y}, \mathbb{Z}(\Delta), \mathscr{U}, \mathscr{Y}, \mathsf{k}),$$

where

(i) U and Y are finite-dimensional $\mathbb{R}$-vector spaces,

(ii) $\mathscr{U}$ is a subspace of $\ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$,

(iii) $\mathscr{Y}$ is a subspace of $\ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$, and

(iv) $\mathsf{k} \colon \mathbb{Z}(\Delta) \to \mathsf{L}(\mathsf{U}; \mathsf{Y})$

Discrete-time convolution systems

Definition (cont'd)

are such that, if we take $\mathsf{K}(t,\tau) = \mathsf{k}(t-\tau)$, then

$$\Sigma' = (\mathsf{U}, \mathsf{Y}, \mathbb{Z}(\Delta), \mathscr{U}, \mathscr{Y}, \mathsf{K})$$

is a summation kernel system.

- We call k a ***convolution kernel***.
- The notion of causality for summation kernel systems transfers easily to discrete-time convolution systems.

Definition

A discrete-time convolution kernel

$$\mathsf{k} \colon \mathbb{Z}(\Delta) \to \mathsf{L}(\mathsf{U}; \mathsf{Y})$$

is ***causal*** if $\mathsf{k}(t) = 0$ for $t \in \mathbb{Z}_{<0}(\Delta)$.

Discrete-time convolution systems

- As with summation kernel systems, one must have conditions on $\mathscr{U}$, $\mathscr{Y}$, and k to ensure continuity of the input/output map.
- One can convert the conditions we have for summation kernel systems, but we have already carefully considered the matter of continuity of convolution on Slide 92. In the causal case, we restrict ourselves to signals that are zero for native time. Thus we have continuity of convolution kernel systems in the following cases:
  1. $\mathscr{U} \subseteq \ell^1(\mathbb{Z}(\Delta); \mathsf{U})$, $\mathscr{Y} \subseteq \ell^1(\mathbb{Z}(\Delta); \mathsf{Y})$, and $\mathsf{k} \in \ell^1(\mathbb{Z}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$;
  2. $\mathscr{U} \subseteq \ell^p(\mathbb{Z}(\Delta); \mathsf{U})$, $\mathscr{Y} \subseteq \ell^q(\mathbb{Z}(\Delta); \mathsf{Y})$, and $\mathsf{k} \in \ell^r(\mathbb{Z}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$, where $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}$;
  3. $\mathscr{U} \subseteq \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{U})$, $\mathscr{Y} \subseteq \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{Y})$, and $\mathsf{k} \in \ell_{\mathrm{loc}}(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$.

Linear discrete-time input/output systems from linear discrete-time state space systems

- On Slide 160 we initiated a programme to produce discrete-time input/output systems from discrete-time state space systems. We saw that there were substantial technical considerations.
- Here we do the same for linear systems, where things are quite a lot easier.
- We shall consider the input/output systems arising from that part of the output equations we had labelled as *term 2*, i.e., the parts coming from the proper impulse transmission map (in the time-varying case) and the proper impulse response (in the constant coefficient case).
- In the continuous-time case, both in the linear and not necessarily linear cases, one must take care of the fact that the input/output map is continuous, and this requires considering various cases of input, output, and system assumptions.

Linear discrete-time input/output systems from linear discrete-time state space systems

- Things are considerable simpler in the discrete-time case, owing to the simpler topological structure of the signal spaces in this case.

Theorem

Let $\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$ *be a linear discrete-time state space system and let* $p \in [1, \infty]$*. Let* $t_0 \in \mathbb{T}$ *and let*

$$\mathscr{U} \subseteq \{\mu \in \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U}) \mid \mu(t) = 0,\ t < t_0\},$$
$$\mathscr{Y} = \{\eta \in \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y}) \mid \eta(t) = 0,\ t < t_0\}.$$

Then

$$\Sigma_{\mathrm{i/o}}(t_0) = (\mathsf{U}, \mathsf{Y}, \mathbb{T}, \mathscr{U}, \mathscr{Y}, \mathsf{pitm}_{\Sigma, t_0})$$

is a causal summation kernel system.

- To state the corresponding result in the constant coefficient case follows along similar lines, adapting the conditions for a causal discrete-time convolution system to be continuous.

Linear discrete-time input/output systems from linear discrete-time state space systems

Theorem

Let $\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{Z}(\Delta), \mathscr{U}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$ *be a linear discrete-time state space system with constant coefficients. Let*

$$\mathscr{U} \subseteq \{\mu \in \ell_{\mathrm{loc}}(\mathbb{Z}(\Delta); \mathsf{U}) \mid \mu(t) = 0,\ t < 0\},$$
$$\mathscr{Y} = \{\eta \in \ell_{\mathrm{loc}}(\mathbb{Z}(\Delta); \mathsf{Y}) \mid \eta(t) = 0,\ t < 0\}.$$

Then

$$\Sigma_{\mathrm{i/o}} = (\mathsf{U}, \mathsf{Y}, \mathbb{Z}(\Delta), \mathscr{U}, \mathscr{Y}, \mathsf{pir}_{\Sigma})$$

is a causal discrete-time convolution system.

- *Punchline:* Summation kernel systems and discrete-time convolutions systems arise in a natural way from linear discrete-time state space systems, and this explains, in part, their importance.

Aside: The generality of integral kernel systems and discrete-time convolution systems

- On Slide 234 we considered when a linear continuous-time input/output system is an integral kernel system or (in the stationary case) a continuous-time convolution system.
- Unlike in the continuous-time case where life was complicated by having to consider distributions as kernels, in the discrete-time case the situation is simple and painless.
- Suppose that we are given a linear continuous-time input/output system with a time-domain $\mathbb{T} \subseteq \mathbb{Z}(\Delta)$, with input and output sets U and Y, and with inputs $\mathscr{U} \subseteq \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{U})$ and $\mathscr{Y} \subseteq \ell_{\mathrm{loc}}(\mathbb{T}; \mathsf{Y})$.
- Let the input/output map be $g \colon \mathscr{U} \to \mathscr{Y}$ and define

$$\mathsf{K}(t, \tau)(u) = g(\tau_\tau^* \mathsf{P} u)(t).$$

- Note that, for $\mu \in \mathscr{U}$, we can write

$$\mu = \sum_{\tau \in \mathbb{T}} \tau_\tau^* \mathsf{P} \mu(\tau).$$

Aside: The generality of integral kernel systems and discrete-time convolution systems

- Then using continuity of g,

$$\begin{aligned} g(\mu)(t) &= g\left(\sum_{\tau \in \mathbb{T}} \tau_\tau^* \mathsf{P} \mu(\tau)\right)(t) = \left(\sum_{\tau \in \mathbb{T}} g(\tau_\tau^* \mathsf{P} \mu(\tau))\right)(t) \\ &= \sum_{\tau \in \mathbb{T}} \mathsf{K}(t, \tau)\mu(\tau) = g_{\mathsf{K}}(\mu)(t). \end{aligned}$$

- Thus the system is a summation kernel system.
- If the system is stationary, one similarly shows that it is a discrete-time convolution system.
- *Punchline:* A *very* large number of linear discrete-time input/output systems are summation kernel systems.
- *Corollary to punchline:* A large number of stationary linear discrete-time input/output systems are discrete-time convolution systems.

Reading for Lecture 31

1. Sections V-6.9.4 and and V-6.9.6.
2. Sections V-6.9.3 and V-6.9.5.

Lecture 32
Continuous-time Laplace transform

- You will have seen the continuous Laplace transform before, under the name "Laplace transform." We will concentrate on some facets of the theory that you may not have seen before.

Definition

For $f \in \mathsf{L}^1_{\mathrm{loc}}(\mathbb{R};\mathbb{F})$ and for $p \in [1,\infty]$, denote

$$I^p(f) = \{x \in \mathbb{R} \mid t \mapsto \mathrm{e}^{-xt} f(t) \text{ is in } \mathsf{L}^p(\mathbb{R};\mathbb{F})\}.$$

We say that f is ***p-Laplace transformable*** if $I^p(f) \neq \varnothing$, and we denote by $\mathsf{LT}^p(\mathbb{R};\mathbb{F})$ the set of p-Laplace transformable functions.

- Here are some basic facts.
  1. If $f \in \mathsf{LT}^p(\mathbb{R};\mathbb{F})$, then $I^p(f)$ is an interval; moreover, any interval is a priori possible.
  2. If $q < p$, then $\mathrm{int}(I^p(f)) \subseteq I^q(f)$; in particular $\mathrm{int}(I^p(f)) \subseteq I^1(f)$ for every p.
  3. If $f \in \mathsf{LT}^p(\mathbb{R};\mathbb{F})$ and if $\inf(\mathrm{supp}(f)) > -\infty$, then $\sup I^p(f) = \infty$.
  4. If $f \in \mathsf{LT}^p(\mathbb{R};\mathbb{F})$ and if $\sup(\mathrm{supp}(f)) < \infty$, then $\inf I^p(f) = -\infty$.

Continuous-time Laplace transform

- We shall be exclusively interested in the cases where f is ***causal*** (meaning that $\inf \mathrm{supp}(f) > -\infty$) or ***strictly causal*** (meaning that $\mathrm{supp}(f) \subseteq \mathbb{R}_{\geq 0}$). We denote
  1. $\alpha^p_{\mathrm{min}}(f) = \inf \mathrm{supp}(I^p(f))$,
  2. $\mathsf{LT}^{p,+}(\mathbb{R};\mathbb{F}) = \{f \in \mathsf{LT}^p(\mathbb{R};\mathbb{F}) \mid f \text{ is causal}\}$, and
  3. $\mathsf{LT}^{p,+}(\mathbb{R}_{\geq 0};\mathbb{F}) = \{f \in \mathsf{LT}^p(\mathbb{R};\mathbb{F}) \mid f \text{ is strictly causal}\}$.
- These $\mathsf{LT}^{p,+}$-spaces will serve as the domain of the continuous Laplace transform.
- We shall also need a codomain.
- For an interval $I \subseteq \mathbb{R}$, denote

$$\mathbb{C}_I = \{z \in \mathbb{C} \mid \ \mathrm{Re}(z) \in I\}$$

  and

$$\mathsf{H}(\mathbb{C}_I;\mathbb{C}) = \{F\colon \mathbb{C}_I \to \mathbb{C} \mid \ F \text{ is holomorphic on } \mathbb{C}_I\}.$$

- For $F \in \mathsf{H}(\mathbb{C}_I;\mathbb{C})$ and $x \in I$, denote

$$\begin{aligned} F_x\colon \mathbb{R} &\to \mathbb{C} \\ y &\mapsto F(x + \mathrm{i}y). \end{aligned}$$

Continuous-time Laplace transform

- Denote

$$\mathsf{H}^p(\mathbb{C}_I;\mathbb{C}) = \{F \in \mathsf{H}(\mathbb{C}_I;\mathbb{C}) \mid \ F_x \in \mathsf{L}^p(\mathbb{R};\mathbb{C}),\ x \in I\}$$

  and

$$\overline{\mathsf{H}}^p(\mathbb{C}_I;\mathbb{C}) = \{F \in \mathsf{H}^p(\mathbb{C}_I;\mathbb{C}) \mid \ \sup\{\|F_x\|_p \mid \ x \in I\} < \infty\}.$$

  The spaces $\overline{\mathsf{H}}^p(\mathbb{C}_{\mathbb{R}_{\geq 0}};\mathbb{C})$ are classical, and are known as ***Hardy spaces***. We will primarily be concerned with the bases $p = \infty$ and $p = 2$.
- Here are some properties of Hardy spaces. For simplicity (and since it is all we care about), we suppose that $\sup(I) = \infty$ and that $-\infty < a = \inf(I)$.
  1. The limit $\lim_{x\to a} F_x$ exists in $\mathsf{L}^p(\mathbb{R};\mathbb{C})$. We will assume that $F|\mathbb{C}_{\{a\}}$ is such that $\lim_{x\to a} F_x = F_a$ (limit in $\mathsf{L}^p(\mathbb{R};\mathbb{C})$).
  2. For "nice" sequences $(z_j)_{j\in\mathbb{Z}_{>0}}$ in $\mathbb{C}_I$ converging to $a + \mathrm{i}y$ ("nice" means they should approach remaining in some cone, i.e., non-tangentially), we have pointwise convergence $\lim_{j\to\infty} F(z_j) = F(a + \mathrm{i}y)$ for almost every y.

Continuous-time Laplace transform

- With the domain and codomain in place, we can define the versions of the continuous Laplace transform we care about.

Definition

If $f \in \mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$, the mapping

$$\begin{aligned}\mathscr{L}_{\mathsf{C}}^{p}(f)\colon \mathbb{C}_{I^1(f)} &\to \mathbb{C}\\ z &\mapsto \int_{\mathbb{R}} f(t)\mathrm{e}^{-zt}\,\mathrm{d}t\end{aligned}$$

is the ***causal continuous p-Laplace transform*** or L^p ***causal CLT*** of f.

- Note that the domain of the function $\mathscr{L}_{\mathsf{C}}^{p}(f)$ is $\mathbb{C}_{I^1(f)}$, no matter the value of p.
- Since $\mathrm{int}(I^p(f)) \subseteq I^1(f)$, the only barrier to having $I^p(f) = I^1(f)$ is that $I^1(f)$ may contain its left endpoint, where as $I^p(f)$ may not.

Continuous-time Laplace transform

- Let us enumerate some of the elementary properties of $\mathscr{L}_{\mathsf{C}}^{p}(f)$.

Proposition

We let $f, g \in \mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$, $a \in \mathbb{C}$, and $s \in \mathbb{R}$.

(i) $\mathscr{L}_{\mathsf{C}}^{p}(f) \in \mathsf{H}(\mathbb{C}_{I^1(f)};\mathbb{C})$*;*

(ii) $af \in \mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$, $\alpha_{\mathsf{min}}^{p}(af) = \alpha_{\mathsf{min}}^{p}(f)$*, and* $\mathscr{L}_{\mathsf{C}}^{p}(af) = a\mathscr{L}_{\mathsf{C}}^{p}(f)$*;*

(iii) $f + g \in \mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$, $\alpha_{\mathsf{min}}^{p}(f+g) \le \max\{\alpha_{\mathsf{min}}^{p}(f), \alpha_{\mathsf{min}}^{p}(g)\}$*, and*

$$\mathscr{L}_{\mathsf{C}}^{p}(f+g)(z) = \mathscr{L}_{\mathsf{C}}^{p}(f)(z) + \mathscr{L}_{\mathsf{C}}^{p}(g)(z), \qquad \mathrm{Re}(z) \in I^1(f) \cap I^1(g);$$

(iv) $\tau_s^* f \in \mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$, $\alpha_{\mathsf{min}}^{p}(\tau_s^* f) = \alpha_{\mathsf{min}}^{p}(f)$*, and*

$$\mathscr{L}_{\mathsf{C}}^{p}(\tau_s^* f)(z) = \mathrm{e}^{-sz}\mathscr{L}_{\mathsf{C}}^{p}(f)(z), \qquad \mathrm{Re}(z) \in I^1(f).$$

- The last property is useful because it allows us to translate some results concerning $\mathsf{LT}^{p,+}(\mathbb{R}_{\ge 0};\mathbb{C})$ to $\mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$.

Continuous-time Laplace transform

- There are some obvious connections between the CLT and the CCFT that are clear when one writes $\mathrm{e}^{-zt} = \mathrm{e}^{-xt}\mathrm{e}^{-\mathrm{i}yt}$ in the integrand of the definition of the CLT. Let us record the correspondences.
  1. If $\mathsf{E}_a(t) = \mathrm{e}^{at}$, we have $\mathscr{L}_{\mathsf{C}}^{p}(f)(\sigma + \mathrm{i}\omega) = \mathscr{F}_{\mathsf{CC}}(f\mathsf{E}_{-\sigma})\left(\frac{\omega}{2\pi}\right)$.
  2. The CLT is injective.
  3. The analogue of the Fourier integral is the ***Fourier–Mellin integral***:

$$\mathsf{FMI}[f](t) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathscr{L}_{\mathsf{C}}^{1}(f)(\sigma + \mathrm{i}\omega)\mathrm{e}^{(\sigma+\mathrm{i}\omega)t}\,\mathrm{d}\omega,$$

where $\sigma \in I^1(f)$. As with the Fourier integral, we disregard the convergence of the integral in the above "definition." Indeed, the study of the convergence of the integral is the study of inversion of the CLT.

Continuous-time Laplace transform

- Results concerning convolution and the continuous Laplace transform are well-known, but are sometimes stated with imprecision.
- A general result for causal signals is the following.

Proposition

Let $p, q, r \in [1, \infty]$ satisfy $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. If $f \in \mathsf{LT}^{p,+}(\mathbb{R};\mathbb{C})$ and $g \in \mathsf{LT}^{q,+}(\mathbb{R};\mathbb{C})$, then $f * g \in \mathsf{LT}^{r,+}(\mathbb{R};\mathbb{C})$, $\alpha^{r}_{\mathsf{min}}(f * g) \le \max\{\alpha^{p}_{\mathsf{min}}(f), \alpha^{q}_{\mathsf{min}}(g)\}$, and

$$\mathscr{L}_{\mathsf{C}}^{r}(f * g)(z) = \mathscr{L}_{\mathsf{C}}^{p}(f)(z)\mathscr{L}_{\mathsf{C}}^{q}(g)(z)$$

for $z \in \mathbb{C}_{(a,\infty)}$, and for any $a > \max\{\alpha^{p}_{\mathsf{min}}(f), \alpha^{q}_{\mathsf{min}}(g), \alpha^{r}_{\mathsf{min}}(f * g)\}$.

- We will see a different result for strictly causal signals later.

Continuous-time Laplace transform

- Let us now concentrate for a moment on strictly causal signals. In this case, the transformed signals have some interesting properties, apart from simply being holomorphic in the region $\mathbb{C}_{I^1(f)}$ where they are defined.

Proposition

Let $p \in [1, \infty]$. *If* $f \in \mathsf{LT}^{p,+}(\mathbb{R}_{\geq 0}; \mathbb{C})$, *then:*

(i) $\mathscr{L}_{\mathsf{C}}^{p}(f) \in \mathsf{C}^0(\mathbb{C}_{I^1(f)}; \mathbb{C})$;

(ii) $\mathscr{L}_{\mathsf{C}}^{p}(f)|\mathbb{C}_{(a,\infty)} \in \mathsf{H}^{\infty}(\mathbb{C}_{(a,\infty)}; \mathbb{C}$ *for every* $a \in I^1(f)$.

- In the case of $p = 2$, something very special happens, rather inline with what happens with the CCFT.

Theorem (L^2-Paley–Wiener Theorem)

$\mathscr{L}_{\mathsf{C}}^{2}$ *is an isomorphism of* $\mathsf{LT}^{2,+}(\mathbb{R}_{\geq 0}; \mathbb{C})$ *with* $\overline{\mathsf{H}}^2(\mathbb{C}_{\mathbb{R}_{\geq 0}}; \mathbb{C})$.

Continuous-time Laplace transform

- Strictly causal continuous Laplace transformable signals also have their own convolution theorem.

Proposition

If $f, g \in \mathsf{LT}^{\infty,+}(\mathbb{R}_{\geq 0}; \mathbb{C})$, *then* $f * g \in \mathsf{LT}^{\infty,+}(\mathbb{R}_{\geq 0}; \mathbb{C})$, $\alpha_{\mathrm{min}}^{\infty}(f * g) \leq \max\{\alpha_{\mathrm{min}}^{\infty}(f), \alpha_{\mathrm{min}}^{\infty}(g)\}$, *and*

$$\mathscr{L}_{\mathsf{C}}^{\infty}(f * g)(z) = \mathscr{L}_{\mathsf{C}}^{\infty}(f)(z)\mathscr{L}_{\mathsf{C}}^{\infty}(g)(z)$$

for $z \in \mathbb{C}_{(a,\infty)}$, *and for any* $a > \max\{\alpha_{\mathrm{min}}^{\infty}(f), \alpha_{\mathrm{min}}^{\infty}(g), \alpha_{\mathrm{min}}^{\infty}(f * g)\}$.

Continuous-time Laplace transform

- The rule for the interaction of the CLT with differentiation must be exercised with care. Here we state the general condition, for higher derivatives.

Proposition

Let $f \in \mathsf{C}^{k-1}(\mathbb{R}_{\geq 0};\mathbb{C})$*, and suppose that* $f^{(a)} \in \mathsf{LT}^{+,\infty}(\mathbb{R}_{\geq 0};\mathbb{C})$*,* $a \in \{0,1,\dots,k-1\}$*, and that* $f^{(k-1)}$ *is locally absolutely continuous with*

$$f^{(k-1)}(t) = f(0+) + \int_0^t f^{(k)}(\tau)\,\mathrm{d}\tau, \qquad t \in \mathbb{R}_{>0},$$

for $f^{(k)} \in \mathsf{L}^1_{\mathsf{loc}}(\mathbb{R}_{\geq 0};\mathbb{C})$*. Then, for* $z \in \mathbb{C}_{\mathrm{int}(I^\infty(f^{(k-1)}))}$*,*

$$\lim_{t\to\infty}\int_0^t f^{(k)}(\tau)\mathrm{e}^{-z\tau}\,\mathrm{d}\tau = z^k\mathscr{L}^\infty_{\mathsf{C}}(f)(z) - f(0+)z^{k-1} - \dots - f^{(k-2)}(0+)z - f^{(k-1)}(0+).$$

- Note that it is *not* a conclusion that $f^{(k)}$ is continuous Laplace transformable; the integral on the left only exists conditionally.

Reading for Lecture 32

1. Section IV-9.1.
2. Section III-7.4.1.

Lecture 33
Discrete-time Laplace transform

- You may have seen the discrete Laplace transform before under the name "z-transform." We will not use this name; it kinda sucks. Again, we will concentrate on some facets of the theory that you may not have seen.
- First of all, we will follow the common practice, not with respect to the name, but in how this transform is defined.
- The "expected" definition of the discrete Laplace transform should be

$$\Delta \sum_{n \in \mathbb{Z}} f(n\delta) \mathrm{e}^{-n\Delta z}.$$

- Instead, we will replace "$\mathrm{e}^{\Delta z}$" with "z," and so the transform becomes

$$\Delta \sum_{n \in \mathbb{Z}} f(n\Delta) z^{-n}.$$

Discrete-time Laplace transform

- We can now make our definition.

Definition

For $f \in \ell_{\mathrm{loc}}(\mathbb{Z}(\Delta); \mathbb{F})$ and for $p \in [1, \infty]$, denote

$$I^p(f) = \{r \in \mathbb{R}_{\geq 0} \mid \ t \mapsto r^{-t/\Delta} f(t) \text{ is in } \ell^p(\mathbb{Z}(\Delta); \mathbb{F})\}.$$

We say that f is ***p-Laplace transformable*** if $I^p(f) \neq \varnothing$, and we denote by $\mathsf{LT}^p(\mathbb{Z}(\Delta); \mathbb{F})$ the set of p-Laplace transformable functions.

- Here are some basic facts.
  1. If $f \in \mathsf{LT}^p(\mathbb{Z}(\Delta); \mathbb{F})$, then $I^p(f)$ is an interval; moreover, any interval is a priori possible.
  2. If $q < p$, then $I^q(f) \subseteq I^p(f)$ and $\mathrm{int}(I^p(f)) \subseteq I^q(f)$; in particular $\mathrm{int}(I^p(f)) \subseteq I^1(f)$ for every p.
  3. If $f \in \mathsf{LT}^p(\mathbb{Z}(\Delta); \mathbb{F})$ and if $\inf(\mathrm{supp}(f)) > -\infty$, then $\sup I^p(f) = \infty$.
  4. If $f \in \mathsf{LT}^p(\mathbb{Z}(\Delta); \mathbb{F})$ and if $\sup(\mathrm{supp}(f)) < \infty$, then $\inf I^p(f) = 0$.

Discrete-time Laplace transform

- We shall be exclusively interested in the cases where f is ***causal*** (meaning that $\inf \operatorname{supp}(f) > -\infty$) or ***strictly causal*** (meaning that $\operatorname{supp}(f) \subseteq \mathbb{Z}_{\geq 0}(\Delta)$). We denote
  1. $\alpha^p_{\mathrm{min}}(f) = \inf \operatorname{supp}(I^p(f))$,
  2. $\mathsf{LT}^{p,+}(\mathbb{Z}(\Delta); \mathbb{F}) = \{f \in \mathsf{LT}^p(\mathbb{Z}(\Delta); \mathbb{F}) \mid f \text{ is causal}\}$, and
  3. $\mathsf{LT}^{p,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{F}) = \{f \in \mathsf{LT}^p(\mathbb{Z}(\Delta); \mathbb{F}) \mid f \text{ is strictly causal}\}$.
- These $\mathsf{LT}^{p,+}$-spaces will serve as the domain of the discrete Laplace transform.
- We shall also need a codomain.
- For an interval $I \subseteq \mathbb{R}_{\geq 0}$, denote

$$\mathbb{A}_I = \{z \in \mathbb{C} \mid |z| \in I\}$$

and

$$\mathsf{H}(\mathbb{A}_I; \mathbb{C}) = \{F \colon \mathbb{A}_I \to \mathbb{C} \mid F \text{ is holomorphic on } \mathbb{A}_I\}.$$

- For $F \in \mathsf{H}(\mathbb{A}_I; \mathbb{C})$ and $r \in I$, denote $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ and

$$\begin{aligned} F_r \colon \mathbb{S}^1 &\to \mathbb{C} \\ (\cos\theta, \sin\theta) &\mapsto F(r\mathrm{e}^{\mathrm{i}\theta}). \end{aligned}$$

Discrete-time Laplace transform

- Denote

$$\mathsf{H}^p(\mathbb{A}_I; \mathbb{C}) = \{F \in \mathsf{H}(\mathbb{A}_I; \mathbb{C}) \mid F_r \in \mathsf{L}^p(\mathbb{S}^1; \mathbb{C}),\ r \in I\}$$

and

$$\overline{\mathsf{H}}^p(\mathbb{A}_I; \mathbb{C}) = \{F \in \mathsf{H}^p(\mathbb{A}_I; \mathbb{C}) \mid \sup\{\|F_r\|_p \mid r \in I\} < \infty\}.$$

The spaces $\overline{\mathsf{H}}^p(\mathbb{A}_{[0,1]}; \mathbb{C})$ are classical, and are known as ***Hardy spaces***. By the conformal transformation $z \mapsto z^{-1}$, these are transformed into $\overline{\mathsf{H}}^p(\mathbb{A}_{[1,\infty)}; \mathbb{C})$, and these are the versions relevant to us. We will primarily be concerned with the bases $p = \infty$ and $p = 2$.

- Here are some properties of Hardy spaces. For simplicity (and since it is all we care about), we suppose that $\sup(I) = \infty$ and that $0 < a = \inf(I)$.
  1. The limit $\lim_{r \to a} F_r$ exists in $\mathsf{L}^p(\mathbb{S}^1; \mathbb{C})$. We will assume that $F|\mathbb{A}_{\{a\}}$ is such that $\lim_{r \to a} F_r = F_a$ (limit in $\mathsf{L}^p(\mathbb{S}^1; \mathbb{C})$).
  2. For "nice" sequences $(z_j)_{j \in \mathbb{Z}_{>0}}$ in $\mathbb{A}_I$ converging to $a\mathrm{e}^{\mathrm{i}\theta}$ ("nice" means they should approach remaining in some cone, i.e., non-tangentially), we have pointwise convergence $\lim_{j \to \infty} F(z_j) = F(a\mathrm{e}^{\mathrm{i}\theta})$ for almost every θ.

Discrete-time Laplace transform

- With the domain and codomain in place, we can define the versions of the discrete Laplace transform we care about.

Definition

If $f \in \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta);\mathbb{C})$, the mapping

$$\begin{aligned}\mathscr{L}_{\mathsf{D}}^{p}(f)\colon \mathbb{A}_{I^1(f)} &\to \mathbb{C}\\ z &\mapsto \Delta \sum_{t\in\mathbb{Z}(\Delta)} f(t)z^{-t/\Delta}\end{aligned}$$

is the ***causal discrete p-Laplace transform*** or ℓ^p ***causal DLT*** of f.

- Note that the domain of the function $\mathscr{L}_{\mathsf{D}}^{p}(f)$ is $\mathbb{A}_{I^1(f)}$, no matter the value of p.
- Since $\operatorname{int}(I^p(f)) \subseteq I^1(f)$, the only barrier to having $I^p(f) = I^1(f)$ is that $I^1(f)$ may contain its left endpoint, where as $I^p(f)$ may not.

Discrete-time Laplace transform

- Let us enumerate some of the elementary properties of $\mathscr{L}_{\mathsf{D}}^{p}(f)$.

Proposition

We let $f, g \in \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta);\mathbb{C})$, $a \in \mathbb{C}$, and $s \in \mathbb{Z}(\Delta)$.

(i) $\mathscr{L}_{\mathsf{D}}^{p}(f) \in \mathsf{H}(\mathbb{A}_{I^1(f)};\mathbb{C})$*;*

(ii) $af \in \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta);\mathbb{C})$, $\alpha_{\mathsf{min}}^{p}(af) = \alpha_{\mathsf{min}}^{p}(f)$*, and* $\mathscr{L}_{\mathsf{D}}^{p}(af) = a\mathscr{L}_{\mathsf{D}}^{p}(f)$*;*

(iii) $f + g \in \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta);\mathbb{C})$, $\alpha_{\mathsf{min}}^{p}(f+g) \le \max\{\alpha_{\mathsf{min}}^{p}(f), \alpha_{\mathsf{min}}^{p}(g)\}$*, and*

$$\mathscr{L}_{\mathsf{D}}^{p}(f+g)(z) = \mathscr{L}_{\mathsf{D}}^{p}(f)(z) + \mathscr{L}_{\mathsf{D}}^{p}(g)(z), \qquad |z| \in I^1(f) \cap I^1(g);$$

(iv) $\tau_s^* f \in \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta);\mathbb{C})$, $\alpha_{\mathsf{min}}^{p}(\tau_s^* f) = \alpha_{\mathsf{min}}^{p}(f)$*, and*

$$\mathscr{L}_{\mathsf{D}}^{p}(\tau_s^* f)(z) = z^{-s/\Delta}\mathscr{L}_{\mathsf{D}}^{p}(f)(z), \qquad |z| \in I^1(f).$$

- The last property is useful because it allows us to translate some results concerning $\mathsf{LT}^{p,+}(\mathbb{Z}_{\ge 0}(\Delta);\mathbb{C})$ to $\mathsf{LT}^{p,+}(\mathbb{Z}(\Delta);\mathbb{C})$.

Discrete-time Laplace transform

- There are some obvious connections between the DLT and the DCFT that are clear when one writes $zt = r^t e^{-i\theta t}$ in the summand of the definition of the DLT. Let us record the correspondences.
  1. If $P_a(j) = a^j$ and $\delta_b(t) = bt$, we have $\mathscr{L}_D^p(f)(re^{i\theta}) = \mathscr{F}_{DC}(f(P_{1/r} \circ \delta_{\Delta^{-1}}))\left(\frac{\theta}{2\pi}\right)$.
  2. The DLT is injective.
  3. The analogue of the Fourier series is

$$f(t) = \frac{r^{t/\Delta}}{2\pi\Delta} \int_0^{2\pi} \mathscr{L}_C^p(f)(re^{i\theta})e^{i\theta t/\Delta} \, d\theta,$$

  where $r \in I^1(f)$. As with the DCFT, this sum converges uniformly and so gives a direct formula for the inverse of the DLT.

Discrete-time Laplace transform

- Results concerning convolution and the discrete Laplace transform are well-known, but are sometimes stated with imprecision.
- A general result for causal signals is the following.

Proposition

Let $p, q, r \in [1, \infty]$ *satisfy* $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. *If* $f \in \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta); \mathbb{C})$ *and* $g \in \mathsf{LT}^{q,+}(\mathbb{Z}(\Delta); \mathbb{C})$, *then* $f * g \in \mathsf{LT}^{r,+}(\mathbb{Z}(\Delta); \mathbb{C})$, $\alpha^r_{\min}(f * g) \leq \max\{\alpha^p_{\min}(f), \alpha^q_{\min}(g)\}$, *and*

$$\mathscr{L}_D^r(f * g)(z) = \mathscr{L}_D^p(f)(z)\mathscr{L}_D^q g)(z)$$

for $z \in \mathbb{A}_{(a,\infty)}$, *and for any* $a > \max\{\alpha^p_{\min}(f), \alpha^q_{\min}(g), \alpha^r_{\min}(f * g)\}$.

- We will see a different result for strictly causal signals later.

Discrete-time Laplace transform

- Let us now concentrate for a moment on strictly causal signals. In this case, the transformed signals have some interesting properties, apart from simply being holomorphic in the region $\mathbb{C}_{I^1(f)}$ where they are defined.

Proposition

Let $p \in [1, \infty]$. *If* $f \in \mathsf{LT}^{p,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{C})$, *then*

(i) $\mathscr{L}^p_{\mathsf{D}}(f) \in \mathsf{C}^0(\mathbb{A}_{I^1(f)}; \mathbb{C})$ *and*

(ii) $\mathscr{L}^p_{\mathsf{D}}(f)|\mathbb{A}_{(a,\infty)} \in \mathsf{H}^\infty(\mathbb{A}_{(a,\infty)}; \mathbb{C})$ *for every* $a \in I^1(f)$.

- In the case of $p = 2$, something very special happens, rather inline with what happens with the CCFT.

Theorem (ℓ^2-Paley–Wiener Theorem)

$\mathscr{L}^2_{\mathsf{D}}$ *is an isomorphism of* $\mathsf{LT}^{2,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{C})$ *with* $\overline{\mathsf{H}}^2(\mathbb{A}_{[1,\infty)}; \mathbb{C})$.

Discrete-time Laplace transform

- Strictly causal discrete Laplace transformable signals also have their own convolution theorem.

Proposition

If $f, g \in \mathsf{LT}^{1,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{C})$, *then* $f * g \in \mathsf{LT}^{1,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathbb{C})$, $\alpha^\infty_{\mathrm{min}}(f * g) \leq \max\{\alpha^1_{\mathrm{min}}(f), \alpha^1_{\mathrm{min}}(g)\}$, *and*

$$\mathscr{L}^1_{\mathsf{D}}(f * g)(z) = \mathscr{L}^1_{\mathsf{D}}(f)(z)\mathscr{L}^1_{\mathsf{D}}(g)(z)$$

for $z \in \mathbb{A}_{(a,\infty)}$, *and for any* $a > \max\{\alpha^1_{\mathrm{min}}(f), \alpha^1_{\mathrm{min}}(g), \alpha^1_{\mathrm{min}}(f * g)\}$.

Discrete-time Laplace transform

- The rule for the interaction of the DLT with differences is analogous to that of the CLT with derivatives. Here we state the general condition, for higher derivatives.

Proposition

If $f \in \mathsf{LT}^{1,+}(\mathbb{Z}_{\geq 0}(\Delta);\mathbb{C})$*, then, for* $z \in \mathrm{int}(\mathbb{A}_{I^1(f)})$*,*

$$\mathscr{L}_{\mathsf{D}}^{1}(\tau^*_{-k\Delta}f)(z) = z^k \mathscr{L}_{\mathsf{D}}^{1}(f)(z) - (\Delta z)^k f(0) - \cdots - (\Delta z)^2 f((k-2)\Delta) - (\Delta z) f((k-1)\Delta).$$

- Note that, unlike in the case of the CLT, here we are allowed to conclude that the forward differences are discrete Laplace transformable.

Reading for Lecture 33

1. Section IV-9.2.
2. Section III-7.5.1.

Lecture 34
Transfer functions for continuous-time convolution systems

- We shall consider continuous-time convolution systems and linear continuous-time state space systems.
- In considering transfer functions, we work with (1) the Laplace transform of the convolution kernel and (2) the Laplace transform of the impulse response.
- There are many systems analysis and design methodologies that have been developed for implementation using the transfer function. We shall not get into this, as this is the realm of subjects like filter design and control theory.
- Instead, we are concerned with the mathematical formulation and properties of transfer functions.

Transfer functions for continuous-time convolution systems

- Our linear system models have all made use of state/input/output spaces that are $\mathbb{R}$-vector spaces. Transfer functions involve the complex variable z. We must "complexify."
  1. If V is a $\mathbb{R}$-vector space, then $\mathsf{V}_{\mathbb{C}}$ is the $\mathbb{C}$-vector space $\mathsf{V}_{\mathbb{C}} = \mathsf{V} \times \mathsf{V}$ with vector space operations
  $$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2), \quad (a + ib)(u, v) = (au - bv, av + bu).$$
  2. If $\mathsf{L} \in \mathsf{L}(\mathsf{U}; \mathsf{V})$ is a $\mathbb{R}$-linear map, we define a $\mathbb{C}$-linear map $\mathsf{L}_{\mathbb{C}} \in \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{V}_{\mathbb{C}})$ by
  $$\mathsf{L}_{\mathbb{C}}(u, v) = (\mathsf{L}(u), \mathsf{L}(v)).$$
- If you are following the path of $\mathsf{X} = \mathbb{R}^n$, $\mathsf{U} = \mathbb{R}^m$, $\mathsf{Y} = \mathbb{R}^k$, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times m}$, $\boldsymbol{C} \in \mathbb{R}^{k\times n}$, and $\boldsymbol{D} \in \mathbb{R}^{k\times m}$ (and it is fine if you are), then you will simply have $\mathsf{X}_{\mathbb{C}} = \mathbb{C}^n$, $\mathsf{U}_{\mathbb{C}} = \mathbb{C}^m$, $\mathsf{Y}_{\mathbb{C}} = \mathbb{C}^k$, $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\boldsymbol{B} \in \mathbb{C}^{n\times m}$, $\boldsymbol{C} \in \mathbb{C}^{k\times n}$, and $\boldsymbol{D} \in \mathbb{C}^{k\times m}$. That is to say, in these cases, complexification is "replace $x \in \mathbb{R}$ with $z \in \mathbb{C}$."

Transfer functions for continuous-time convolution systems

- We can now define the transfer function for continuous-time convolution systems.

Definition

Let $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathscr{U}, \mathscr{Y}, \mathsf{k})$ be a continuous-time convolution system and suppose that $\mathsf{k} \in \mathsf{LT}^{+,p}(\mathbb{R}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$. The ***transfer function*** for Σ is the mapping

$$\begin{aligned} \mathsf{T}_\Sigma \colon \mathbb{C}_{I^1(\mathsf{k})} &\to \mathsf{L}(\mathsf{U}_\mathbb{C}; \mathsf{Y}_\mathbb{C}) \\ z &\mapsto \mathscr{L}^p_{\mathsf{C}}(\mathsf{k})(z). \end{aligned}$$

- The idea of the transfer function approach is the following transformation rule for the input/output map upon taking Laplace transforms:

$$g_\mathsf{k}(\mu)(t) = \int_\mathbb{R} \mathsf{k}(t-\tau)(\mu(\tau))\,\mathrm{d}\tau \quad \Longrightarrow \quad \mathscr{L}_\mathsf{C}(g_\mathsf{k}(\mu))(z) = T_\Sigma(z)\mathscr{L}_\mathsf{C}(\mu)(z).$$

Transfer functions for continuous-time convolution systems

- However, any assertion such as this requires conditions on inputs, outputs, and the convolution kernel.
- Here are two kinds of conditions, obviously simply derived from the relationships we have seen for convolution and the Laplace transform.
  1. for $p, q, r \in [1, \infty]$ satisfying $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}$,
     (a) $\mathsf{k} \in \mathsf{LT}^{r,+}(\mathbb{R}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
     (b) $\mathscr{U} \subseteq \mathsf{LT}^{p,+}(\mathbb{R}; \mathsf{U})$, and
     (c) $\mathscr{Y} \subseteq \mathsf{LT}^{q,+}(\mathbb{R}; \mathsf{Y})$;
  2. (a) $\mathsf{k} \in \mathsf{LT}^{\infty,+}(\mathbb{R}_{\geq 0}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
     (b) $\mathscr{U} \subseteq \mathsf{LT}^{\infty,+}(\mathbb{R}_{\geq 0}; \mathsf{U})$, and
     (c) $\mathscr{Y} \subseteq \mathsf{LT}^{\infty,+}(\mathbb{R}_{\geq 0}; \mathsf{Y})$.
- Another useful condition is rather connected to the L^2-Paley–Wiener Theorem from Slide 280.
  3. (a) $\mathsf{k} \in \mathsf{L}^1(\mathbb{R}_{\geq 0}; \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
     (b) $\mathscr{U} \subseteq \mathsf{L}^2(\mathbb{R}_{\geq 0}; \mathsf{U})$, and
     (c) $\mathscr{Y} \subseteq \mathsf{L}^2(\mathbb{R}_{\geq 0}; \mathsf{U})$.

  In this case we also get a bound of the H^2-norms of the input/output map:

$$\|\mathscr{L}^2_\mathsf{C}(\mathsf{k} * \mu)\|_{\mathsf{H}^2,\mathbb{R}_{\geq 0}} \leq \|\mathscr{L}^1_\mathsf{C}(\mathsf{k})\|_{\mathsf{H}^\infty,\mathbb{R}_{\geq 0}} \|\mathscr{L}^2_\mathsf{C}(\mu)\|_{\mathsf{H}^2,\mathbb{R}_{\geq 0}}.$$

Transfer functions for linear continuous-time state space systems

- We now consider linear continuous-time state space systems.

Definition

For a linear continuous-time state space system

$$\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{R}, \mathscr{U}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$$

with constant coefficients, the ***transfer function*** is the $\mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}})$-valued function

$$\begin{aligned} \mathsf{T}_\Sigma \colon \mathbb{C}_{(\sigma_{\mathsf{max}}(\mathsf{A}),\infty)} &\to \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}}) \\ z &\mapsto \mathsf{C}_{\mathbb{C}} \circ (z\,\mathrm{id}_{\mathsf{X}_{\mathbb{C}}} - \mathsf{A}_{\mathbb{C}})^{-1} \circ \mathsf{B}_{\mathbb{C}} + \mathsf{D}_{\mathbb{C}}, \end{aligned}$$

where

$$\sigma_{\mathsf{max}}(\mathsf{A}) = \max\{\mathrm{Re}(\lambda) \mid \lambda \in \mathrm{spec}(\mathsf{A})\}.$$

Transfer functions for linear continuous-time state space systems

- This transfer function enjoys the following properties:
  1. $T_\Sigma = \mathscr{L}^1_{\mathsf{C}}(\mathrm{ir}_\Sigma)$;
  2. upon choosing bases for X, U, and Y, T_Σ is a matrix whose entries are proper (strictly proper, if $\mathsf{D} = 0$) rational functions;
  3. for any $a > \sigma_{\mathsf{max}}(\mathsf{A})$, $\mathsf{T}_\Sigma|\mathbb{C}_{[a,\infty)} \in \mathsf{H}^\infty(\mathbb{C}_{[a,\infty)}; \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}}))$;
  4. if $\mathsf{D} = 0$, then, for any $a > \sigma_{\mathsf{max}}(\mathsf{A})$, $\mathsf{T}_\Sigma|\mathbb{C}_{[a,\infty)} \in \mathsf{H}^2(\mathbb{C}_{[a,\infty)}; \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}}))$.
- Consider especially the case where $\sigma_{\mathsf{max}}(\mathsf{A}) < 0$.

Transfer functions for discrete-time convolution systems

- We proceed very much as we did in the continuous-time case.

Definition

Let $\Sigma = (\mathsf{U}, \mathsf{Y}, \mathscr{U}, \mathscr{Y}, \mathsf{k})$ be a discrete-time convolution system and suppose that $\mathsf{k} \in \mathsf{LT}^{+,p}(\mathbb{Z}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$. The ***transfer function*** for Σ is the mapping

$$\begin{aligned} \mathsf{T}_\Sigma \colon \mathbb{A}_{I^1(\mathsf{k})} &\to \mathsf{L}(\mathsf{U}_\mathbb{C}; \mathsf{Y}_\mathbb{C}) \\ z &\mapsto \mathscr{L}^p_\mathsf{D}(\mathsf{k})(z). \end{aligned}$$

- The idea of the transfer function approach is the following transformation rule for the input/output map upon taking Laplace transforms:

$$g_\mathsf{k}(\mu)(k\Delta) = \sum_{j\in\mathbb{Z}} \mathsf{k}((k-j)\Delta)(\mu(j\Delta\tau)) \quad \Longrightarrow \quad \mathscr{L}_\mathsf{D}(g_\mathsf{k}(\mu))(z) = T_\Sigma(z)\mathscr{L}_\mathsf{D}(\mu)(z).$$

Transfer functions for discrete-time convolution systems

- However, any assertion such as this requires conditions on inputs, outputs, and the convolution kernel.
- Here are two kinds of conditions, obviously simply derived from the relationships we have seen for convolution and the Laplace transform.
  1. for $p, q, r \in [1, \infty]$ satisfying $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}$,
     (a) $\mathsf{k} \in \mathsf{LT}^{r,+}(\mathbb{Z}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
     (b) $\mathscr{U} \subseteq \mathsf{LT}^{p,+}(\mathbb{Z}(\Delta); \mathsf{U})$, and
     (c) $\mathscr{Y} \subseteq \mathsf{LT}^{q,+}(\mathbb{Z}(\Delta); \mathsf{Y})$;
  2. (a) $\mathsf{k} \in \mathsf{LT}^{\infty,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
     (b) $\mathscr{U} \subseteq \mathsf{LT}^{\infty,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{U})$, and
     (c) $\mathscr{Y} \subseteq \mathsf{LT}^{\infty,+}(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{Y})$.
- Another useful condition is rather connected to the ℓ^2-Paley–Wiener Theorem from Slide 292.
  3. (a) $\mathsf{k} \in \ell^1(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{L}(\mathsf{U}; \mathsf{Y}))$,
     (b) $\mathscr{U} \subseteq \ell^2(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{U})$, and
     (c) $\mathscr{Y} \subseteq \ell^2(\mathbb{Z}_{\geq 0}(\Delta); \mathsf{U})$.

  In this case we also get a bound of the H^2-norms of the input/output map:

$$\|\mathscr{L}^2_\mathsf{D}(\mathsf{k} * \mu)\|_{\mathsf{H}^2,[1,\infty)} \leq \|\mathscr{L}^1_\mathsf{D}(\mathsf{k})\|_{\mathsf{H}^\infty,[1,\infty)} \|\mathscr{L}^2_\mathsf{D}(\mu)\|_{\mathsf{H}^2,[1,\infty)}.$$

Transfer functions for linear discrete-time state space systems

- Again, we follow closely the continuous-time case. Note carefully the difference in the relationship between the eigenvalues and the region of definition of the transfer function.

Definition

For a linear discrete-time state space system

$$\Sigma = (\mathsf{X}, \mathsf{U}, \mathsf{Y}, \mathbb{Z}(\Delta), \mathscr{U}, \mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D})$$

with constant coefficients, the ***transfer function*** is the $\mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}})$-valued function

$$\begin{aligned} \mathsf{T}_\Sigma \colon \mathbb{A}_{(\rho_{\max}(\mathsf{A}),\infty)} &\to \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}}) \\ z &\mapsto \mathsf{C}_{\mathbb{C}} \circ (z\,\mathrm{id}_{\mathsf{X}_{\mathbb{C}}} - \mathsf{A}_{\mathbb{C}})^{-1} \circ \mathsf{B}_{\mathbb{C}} + \mathsf{D}_{\mathbb{C}}, \end{aligned}$$

where

$$\rho_{\max}(\mathsf{A}) = \max\{|\lambda| \mid \lambda \in \mathrm{spec}(\mathsf{A})\}.$$

Transfer functions for linear discrete-time state space systems

- This transfer function enjoys the following properties:
  1. $T_\Sigma = \mathscr{L}^1_{\mathsf{D}}(\mathrm{ir}_\Sigma)$;
  2. upon choosing bases for X, U, and Y, T_Σ is a matrix whose entries are proper (strictly proper, if $\mathsf{D} = 0$) rational functions;
  3. for any $a > \rho_{\max}(\mathsf{A})$, $\mathsf{T}_\Sigma|\mathbb{A}_{[a,\infty)} \in \mathsf{H}^\infty(\mathbb{A}_{[a,\infty)}; \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}}))$;
  4. if $\mathsf{D} = 0$, then, for any $a > \rho_{\max}(\mathsf{A})$, $\mathsf{T}_\Sigma|\mathbb{A}_{[a,\infty)} \in \mathsf{H}^2(\mathbb{A}_{[a,\infty)}; \mathsf{L}(\mathsf{U}_{\mathbb{C}}; \mathsf{Y}_{\mathbb{C}}))$.
- Consider especially the case where $\rho_{\max}(\mathsf{A}) < 1$.

Reading for Lecture 34

1. Sections V-7.1.1 and V-7.2.1.
2. Sections V-7.1.2 and V-7.2.2.
3. Sections V-7.1.4 and V-7.2.4.