

European Embedded Control Institute

Graduate School on Control — Spring 2010

The Behavioral Approach to Modeling and Control

EXERCISES

Preamble

We collect a few facts about polynomials and rational functions that come in useful in some of the exercises.

Dividing with rest

For $a, b \in \mathbb{R}[\xi]$, $b \neq 0$, there exist $d, r \in \mathbb{R}[\xi]$ with $\text{degree}(r) < \text{degree}(b)$ or $r = 0$, such that

$$a = bd + r.$$

The proof is easy.

gcd and lcm

Let $a, b \in \mathbb{R}[\xi]$. The *greatest common divisor* of a, b (denoted $\text{gcd}(a, b)$) is the monic polynomial of largest degree that is a factor of both a and b . The *least common multiple* of a, b (denoted $\text{lcm}(a, b)$) is the monic polynomial of smallest degree that has both a and b as a factor.

If $\text{gcd}(a, b) = 1$, then a and b are said to be *coprime*.

The Bézout identity

Let $a, b \in \mathbb{R}[\xi]$. Then a and b are coprime if and only if there exist $x, y \in \mathbb{R}[\xi]$ such that

$$ax + by = 1. \quad (\text{Bézout identity})$$



Étienne Bézout (1730-1783)

Proof: The (*if*) part is easy. The (*only if*) part can be shown as follows. Assume that a and b are coprime. Let $\ell \in \mathbb{R}[\xi]$ be the monic polynomial of least degree that can be written as $ax + by$ with $x, y \in \mathbb{R}[\xi]$. We need to prove that $\ell = 1$. Division with rest gives $a = \ell d + r$ with $\text{degree}(r) < \text{degree}(\ell)$, or $r = 0$. Let $\ell = a\zeta + b\eta$. Then $r = a(1 - \zeta d) - b\eta d$. Hence r can also be written as $ax + by$. Since r cannot be nonzero and have $\text{degree}(r) < \text{degree}(\ell)$, we conclude that $r = 0$. Hence ℓ divides a . Similarly, it follows that ℓ divides b . Hence, since a and b are coprime, $\ell = 1$. This yields the Bézout identity.

The Hankel matrix

Let $g \in \mathbb{R}(\xi)$ be strictly proper. Consider its ‘Laurent series’ expansion

$$g(\xi) = \frac{g_1}{\xi} + \frac{g_2}{\xi^2} + \dots + \frac{g_n}{\xi^n} + \dots$$

The *Hankel matrix* associated with g is defined as

$$\mathbb{H}(g) := \begin{bmatrix} g_1 & g_2 & \dots & g_{n-1} & g_n & \dots \\ g_2 & g_3 & \dots & g_n & g_{n+1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{n-1} & g_n & \dots & g_{2n-3} & g_{2n-2} & \dots \\ g_n & g_{n+1} & \dots & g_{2n-2} & g_{2n-1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Matrices (infinite as well as finite) which have the same entries parallel to the ‘reverse’ diagonal are called *Hankel matrices*. In a picture,

Hankel \rightsquigarrow
$$\begin{bmatrix} \cdot & \cdot & \cdot & \clubsuit & \heartsuit & \diamond & \spadesuit & \dots \\ \cdot & \cdot & \clubsuit & \heartsuit & \diamond & \spadesuit & \cdot & \dots \\ \cdot & \clubsuit & \heartsuit & \diamond & \spadesuit & \cdot & \cdot & \dots \\ \clubsuit & \heartsuit & \diamond & \spadesuit & \cdot & \cdot & \cdot & \dots \\ \heartsuit & \diamond & \spadesuit & \cdot & \cdot & \cdot & \cdot & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$



Hermann Hankel (1839-1873)

Matrices (infinite as well as finite) which have the same entries parallel to the diagonal are called *Toeplitz matrices*. In a picture,

Toeplitz \rightsquigarrow
$$\begin{bmatrix} \heartsuit & \diamond & \spadesuit & \cdot & \cdot & \cdot & \cdot & \dots \\ \clubsuit & \heartsuit & \diamond & \spadesuit & \cdot & \cdot & \cdot & \dots \\ \cdot & \clubsuit & \heartsuit & \diamond & \spadesuit & \cdot & \cdot & \dots \\ \cdot & \cdot & \clubsuit & \heartsuit & \diamond & \spadesuit & \cdot & \dots \\ \cdot & \cdot & \cdot & \clubsuit & \heartsuit & \diamond & \spadesuit & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$



Otto Toeplitz (1881-1940)

The *rank* of an infinite matrix is defined as the supremum of the dimension of its nonsingular submatrices. The following result determines the rank of $\mathbb{H}(g)$ in terms of the degree of the denominator of g .

$\text{rank}(\mathbb{H}(g))$ is finite and equal to the degree of the polynomial a of lowest degree such that ag is polynomial.

Proof: Let $r = \text{rank}(\mathbb{H}(g))$ and $n =$ the degree of the polynomial $a \in \mathbb{R}[\xi]$ of lowest degree such that ag is polynomial. Denote the columns of $\mathbb{H}(g)$ by $H_1, H_2, \dots, H_k, \dots$

(i) Assume that ag is polynomial with $a \in \mathbb{R}[\xi]$ of degree n . Write $a(\xi)g(\xi)$ in terms of its powers of ξ . Since ag is polynomial, the non-positive powers have zero coefficients. This implies that the H_{n+1} is a linear combination of H_1, H_2, \dots, H_n . By the Hankel structure of $\mathbb{H}(g)$ this implies that the H_{n+k} is a linear combination of H_1, H_2, \dots, H_n for all $k = 1, 2, \dots$. Hence $r = \text{rank}(\mathbb{H}(g)) = \text{rank}([H_1 \ H_2 \ \dots \ H_n]) \leq n$.

(ii) From the Hankel structure of $\mathbb{H}(g)$, it follows that if H_{k+1} is a linear combination of H_1, H_2, \dots, H_k , then $H_{k+k'}$ is also a linear combination of H_1, H_2, \dots, H_k for all $k' = 2, 3, \dots$. Therefore $r = \text{rank}(\mathbb{H}(g))$ implies that $r = \text{rank}([H_1 \ H_2 \ \dots \ H_r])$ and that H_{r+1} is a linear combination of H_1, H_2, \dots, H_r . By this linear dependence, there exist $a_0, a_1, \dots, a_r \in \mathbb{R}$ such that $a_0 H_1 + a_1 H_2 + \dots + a_{r-1} H_r + H_{r+1} = 0$. Now conclude that $(a_0 + a_1 \xi + \dots + a_{r-1} \xi^{r-1} + \xi^r)g(\xi)$ is polynomial. Therefore $n \leq r$.

Exercise 1: Linear models

The mathematical model $(\mathbb{R}^n, \mathcal{B})$ is said to be

linear : \Leftrightarrow $[\mathcal{B}$ is a linear subspace of $\mathbb{R}^n]$.

1. Prove that a linear behavior admits a representation

$$Rw = 0, \quad R \in \mathbb{R}^{\bullet \times n}.$$

Call this representation a **kernel representation** of \mathcal{B} , and a **minimal** one if, among all kernel representations of \mathcal{B} , $\text{rowdimension}(R)$ is as small as possible.

2. Prove that $Rw = 0$ is minimal if and only if R has full row rank.
3. How are the R 's corresponding to minimal kernel representations of \mathcal{B} related?
4. Define what you mean by an image representation. Prove its existence.

Exercise 2: Codes

As mentioned in Lecture I, our concepts of ‘universum’ and ‘behavior’ have block codes as an interesting example.

An important class of block codes are subsets of $GF(2)$, that is $\mathcal{B} \subseteq \{0, 1\}^n$; n is called the ‘length’ of the block code. As an example, take $n = 5$, and

$$\mathcal{B} = \{00000, 10110, 01011, 11101\}.$$

1. How many words are there in \mathcal{U} and in \mathcal{B} ?
2. Is \mathcal{B} linear?

In coding applications it is important, for error detection and correction, that code words are as far away from each other as possible. The separation between two code words can be measured by the ‘Hamming distance’, defined by

$$\begin{aligned} d(w_1 w_2 \cdots w_n, w'_1 w'_2 \cdots w'_n) \\ = \text{the number of indices } k \text{ such that } w_k \neq w'_k. \end{aligned}$$

The Hamming distance of a code is defined as the minimum of the distances between unequal code words.

3. Compute the Hamming distance of \mathcal{B} .
4. Prove that this code can correct one error, in the sense that if the codeword is transmitted and received with at most one error, then this error can be corrected.
5. Prove that this code can detect up to two errors, in the sense that if the codeword is transmitted and received with at most two errors, then this can be detected.

Exercise 3: Parity check

Consider the DES with $\mathcal{U} = \{0, 1\}^{32}$ and

$$\mathcal{B} = \left\{ a_1 a_2 \cdots a_{31} a_{32} \mid a_k \in \{0, 1\} \text{ and } a_{32} \equiv \sum_{k=1}^{31} a_k \pmod{2} \right\},$$

the set of 32-bit strings with a parity check as last bit.

1. In what sense is this a linear model?
2. Give a kernel representation of this behavior.
3. Give an image representation of this behavior.
4. Call $e = \sum_{k=1}^{32} a_k$ the **syndrome** associated with this 32-bit string, and explain how e can be used for error detection.
How many errors can this code detect?

Exercise 4: Symmetry

A **transformation group** on a set A is a set of maps that form a subgroup of the group of bijections on A . In other words, there is a group \mathcal{G} and a map T from \mathcal{G} to the bijections on A , such that for all $g, g_1, g_2 \in \mathcal{G}$, there holds

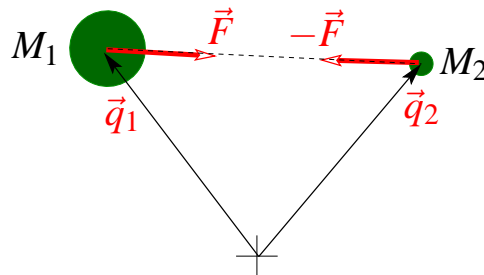
1. $T_1 = \text{id}_A$ (id_A denotes the identity map on A),
2. $T_{g^{-1}} = T_g^{-1}$,
3. $T_{g_1 g_2} = T_{g_2} \circ T_{g_1}$ (\circ denotes map composition).

Let $T_{\mathcal{G}}$ be a transformation group on \mathcal{U} . The mathematical model $(\mathcal{U}, \mathcal{B})$ is said to be **symmetric** with respect to $T_{\mathcal{G}}$ if

$$T_g(\mathcal{B}) = \mathcal{B} \quad \text{for all } g \in \mathcal{G}.$$

1. Identify an obvious symmetry for the 32-bit strings with a parity check discussed in Exercise 3.
2. Formalize time-invariance as a symmetry.
3. Identify a few symmetries for the gravitational attraction of two bodies, $\mathcal{U} = \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$;

$$\mathcal{B} = \left\{ (\vec{q}_1, \vec{q}_2, \vec{F}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \mid \vec{F} = M_1 M_2 \frac{(\vec{q}_2 - \vec{q}_1)}{|\vec{q}_1 - \vec{q}_2|^3} \right\}. \text{ Extend the def-}$$



inition of \mathcal{B} by including M_1, M_2 , so that exchanging the masses also becomes a symmetry.

4. Explain in what sense Maxwell's equations are symmetric with respect to space translation and rotation.

Exercise 5: Memoryless systems

The dynamical system $(\mathbb{T}, \mathbb{W}, \mathcal{B})$ is said to be **memoryless** $:\Leftrightarrow$

$$\llbracket w_1, w_2 \in \mathcal{B} \text{ and } t' \in \mathbb{T} \rrbracket \Rightarrow \llbracket w_1 \wedge_{t'} w_2 \in \mathcal{B} \rrbracket,$$

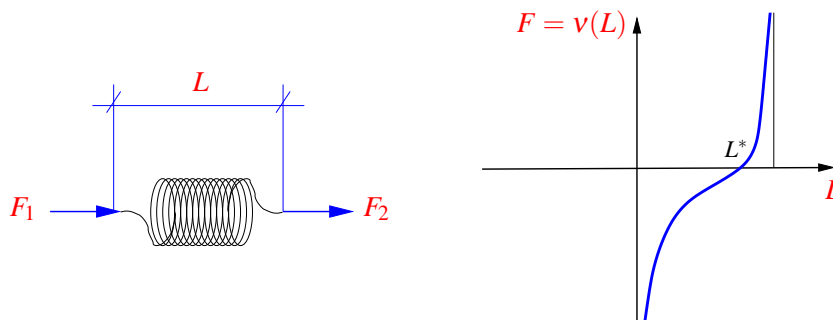
where $w_1 \wedge_{t'} w_2$, the **concatenation** of w_1 and w_2 at t' , is defined as

$$(w_1 \wedge_{t'} w_2)(t) = \begin{cases} w_1(t) & \text{for } t < t', \\ w_2(t) & \text{for } t \geq t'. \end{cases}$$

1. Which of the following physical devices discussed in Lecture I define memoryless systems?

- ▶ The gas law.
- ▶ A resistor, an inductor, a capacitor.
- ▶ The gravitational attraction of two bodies, Kepler's laws, Newton's second law.

Consider a simple spring.



Consider a real variable, E , the **energy** stored in the spring, related to L by

$$E(L) = \int_{L^*}^L v(\sigma) d\sigma,$$

with $-F_1 = F_2 = F = v(L)$ the spring characteristic and L^* the equilibrium length (corresponding to $F = 0$).

2. Consider F, L, E as functions of time. Prove that $\frac{d}{dt}E = F \frac{d}{dt}L$.

3. Prove that the spring viewed in terms of the variables (F, L) defines a memoryless system, but that in terms of the variables (F, L, E) it does not define a memoryless system.

Energy and power considerations can hence bring in dynamics, even in an otherwise memoryless system.

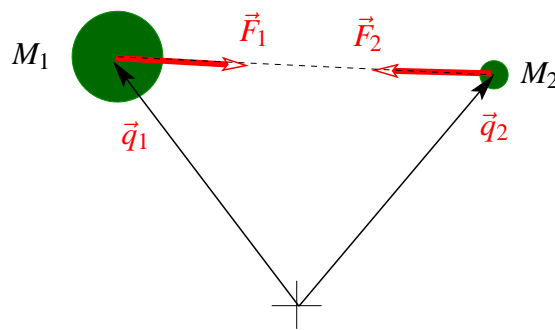
Exercise 6: Two-body problem

The motion of a pointmass in a force field is governed by

$$M \frac{d^2}{dt^2} \vec{q} = \vec{F}(\vec{q}),$$

with M the mass of the body, \vec{q} the position, and $\vec{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ the force field. Studying the resulting motions of the pointmass is called the ‘one-body problem’.

Consider the motion of two bodies, under mutual attraction.



Newton’s second law yields the equations of motion

$$M_1 \frac{d^2}{dt^2} \vec{q}_1 = \vec{F}_1, \quad M_2 \frac{d^2}{dt^2} \vec{q}_2 = \vec{F}_2. \quad (\diamond)$$

According to Newton’s third law,

$$\vec{F}_1 + \vec{F}_2 = \vec{0}. \quad (\spadesuit)$$

The problem is to obtain an explicit description of the trajectories $\vec{q}_1 : \mathbb{R} \rightarrow \mathbb{R}^3, \vec{q}_2 : \mathbb{R} \rightarrow \mathbb{R}^3$ that are possible. This is the so-called ‘two-body problem’, which, in contrast to the three-body or n-body problem, can be reduced to one-body problems.

1. Define the ‘barycenter’ of the two bodies

$$\vec{R} = \frac{M_1 \vec{q}_1 + M_2 \vec{q}_2}{M_1 + M_2}. \quad (\clubsuit)$$

Eliminate $\vec{q}_1, \vec{q}_2, \vec{F}_1, \vec{F}_2$ from $(\diamond, \spadesuit, \clubsuit)$ and show that the behavior of \vec{R} is governed by $\frac{d^2}{dt^2} \vec{R} = 0$.

2. Consider the difference vector of the bodies

$$\vec{\Delta} = \vec{q}_1 - \vec{q}_2. \quad (\heartsuit)$$

Assume that \vec{F}_1 is a function of (\vec{q}_1, \vec{q}_2) only through $\vec{\Delta}$ (makes perfect sense physically). Eliminate $\vec{q}_1, \vec{q}_2, \vec{F}_2$ from $(\diamond, \spadesuit, \heartsuit)$ and prove that the behavior of $(\vec{\Delta}, \vec{F}_1)$ is governed by $\mu \frac{d^2}{dt^2} \vec{\Delta} = \vec{F}_1$, with $\mu = \frac{M_1 M_2}{M_1 + M_2}$; μ is called the ‘reduced mass’. Prove that the motion of $\vec{\Delta}$ is that of one body with mass μ under the force field $\vec{F}_1(\vec{\Delta})$. Hence after solving two one-body problems, we obtain (\vec{q}_1, \vec{q}_2) by

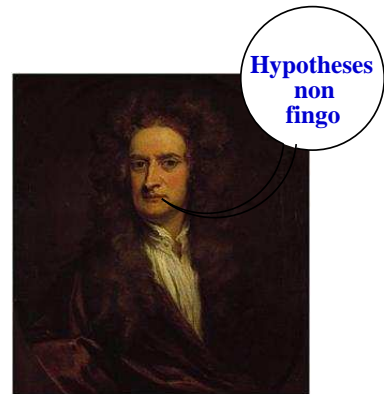
$$\vec{q}_1 = \vec{R} + \frac{M_2}{M_1 + M_2} \vec{\Delta}, \quad \vec{q}_2 = \vec{R} + \frac{M_1}{M_1 + M_2} \vec{\Delta}.$$

3. Often \vec{F}_1 is a central force, that is, it is of the form $\vec{F}_1 = F(\|\vec{\Delta}\|) \frac{\vec{\Delta}}{\|\vec{\Delta}\|}$.

A special case is ‘Kepler’s problem’, with $F(\|\vec{\Delta}\|) = -\frac{1}{\|\vec{\Delta}\|^2}$ (with suitable units). This yields

$$\frac{d^2}{dt^2} \vec{\Delta} + \frac{1}{\|\vec{\Delta}\|^2} \vec{\Delta} = 0 \quad (\star)$$

as the equation for $\vec{\Delta}$. It can be shown that the orbits satisfying K1, K2, K3 (with suitably chosen constants) are solutions. Actually Newton *derived* (\star) from K1, K2, K3.



Isaac Newton (1643-1727)

Do K1, K2, K3 give all the solutions to (\star) ? Argue from physical insight, do not attempt to answer using mathematical arguments.

Exercise 7: The MPUM

Assume that the set of system trajectories $\mathbb{D} = \{\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_n\}$, with $\tilde{w}_k : \mathbb{T} \rightarrow \mathbb{W}$, for $k = 1, 2, \dots, n$, is observed. The following is a version of the (deterministic) **system identification** problem.

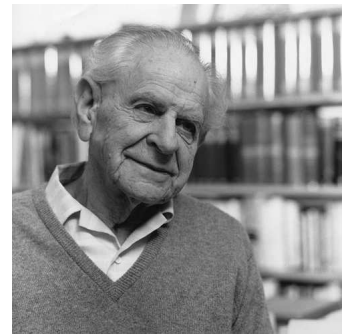
Find the behavior of the dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathcal{B})$ that produced these observations.

Call \mathcal{B} [*unfalsified* by \mathbb{D}] $:\Leftrightarrow [\mathbb{D} \subseteq \mathcal{B}]$.

Call [\mathcal{B}_1 *more powerful* than \mathcal{B}_2] $:\Leftrightarrow [\mathcal{B}_1 \subset \mathcal{B}_2]$.

The more a model forbids, the better it is.

According to Popper, this is against common belief.



Karl Popper (1902-1994)

Let \mathbb{B} be a set of behaviors, i.e., a set of subsets of $W^{\mathbb{T}}$.

Call \mathcal{B}^* [the **most powerful unfalsified model** (MPUM) in \mathbb{B} for \mathbb{D}] $:\Leftrightarrow [$

- ▶ $\mathcal{B}^* \in \mathbb{B}$,
- ▶ \mathcal{B}^* is unfalsified by \mathbb{D} ,
- ▶ \mathcal{B}^* is more powerful than every other element of \mathbb{B} that is unfalsified by \mathbb{D} .]

1. Prove that when $\tilde{w}_k : \mathbb{Z} \rightarrow \mathbb{R}^w$ for $k = 1, 2, \dots, n$, there exists an MPUM in the class of discrete-time LTIDSs.

2. With the result of Exercise 7, part 3, you may also prove that when $\tilde{w}_k \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^w)$ for $k = 1, 2, \dots, n$, there exists an MPUM in \mathcal{L}^w .

Exercise 8: PDEs

A specification of the behavior in terms of an ODE or a PDE is often not very helpful, and there exist other specifications that give much more insight in the nature of \mathcal{B} . For example, Kepler's laws give much more insight than the associated ODE (equation (\star) in Exercise 6).

1. Consider the wave equation

$$\frac{\partial^2}{\partial t^2} w = \frac{\partial^2}{\partial x^2} w.$$

It defines a system $(\mathbb{R}^2, \mathbb{R}, \mathcal{B})$. Prove that

$$\mathcal{B} = \left\{ w \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R}) \mid \right. \\ \left. \exists f_-, f_+ \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \text{ such that } w(t, x) = f_-(t - x) + f_+(t + x) \right\}.$$

Argue that this description of \mathcal{B} is more insightful than the PDE and puts in evidence the wave nature of the behavior.

2. Write Maxwell's equations in polynomial matrix form

$$R \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right) w = 0.$$

Specify the associated set of independent variables, of dependent variables, the w , and the polynomial matrix R .

Exercise 9: The Smith form

Prove the Smith canonical form.

Proceed as follows.

- 1. Assume $M \neq 0$. Prove that by pre- and postmultiplying by a permutation matrix, we may assume that the $(1, 1)$ element of M is $\neq 0$ and has the least degree of all other nonzero elements of M .**
- 2. Let $M_{1,1}$ be this $(1, 1)$ element. Assume there is another nonzero element in the first row or the first column of M . Call this element x . Use division with rest to define r by**

$$x = M_{1,1}d + r \text{ with } r = 0 \text{ or } \mathbf{degree}(r) < \mathbf{degree}(M_{1,1}).$$

Prove that there exist a unimodular matrix U or V such that UM or MV has either one more zero element in the first row or column than M , or a $(1, 1)$ element with degree less than the degree of $M_{1,1}$.

- 3. Prove that in a finite number of steps this leads to a matrix of the form**

$$\begin{bmatrix} M_{1,1} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{bmatrix}.$$

- 4. Obtain the Smith form by induction.**

Exercise 10: Discrete-time systems

1. Prove that for $\mathbb{T} = \mathbb{Z}$, a LITDS has both a kernel representation $R(\sigma)w = 0$ and $R(\sigma^{-1})w = 0$ with $\mathbb{R} \in \mathbb{R}[\xi]^{\bullet \times \bullet}$. Is this also true for $\mathbb{T} = \mathbb{Z}_+$?
2. Give an example of a difference equation $R(\sigma)w = 0$ so that the solution set for $\mathbb{T} = \mathbb{Z}_+$ is not the solution set for $\mathbb{T} = \mathbb{Z}$ restricted to \mathbb{Z}_+ .
3. Formulate the analogue of Proposition 1 for discrete-time systems with $\mathbb{T} = \mathbb{Z}$ and with $\mathbb{T} = \mathbb{Z}_+$.
4. Consider the convolutional code discussed in Lecture I. It is a linear time-invariant difference system. Identify the associated time-set, signal space, and polynomial matrix.

Exercise 11: Minimal representations of LTIDSs

1. Give a formal definition of the greatest common divisor (gcd) and least common multiple (lcm) of $p_1, p_2, \dots, p_n \in \mathbb{R}[\xi]$

2. Let $\mathcal{B}_1, \mathcal{B}_2 \in \mathcal{L}^1$ have kernel representations

$$p_1\left(\frac{d}{dt}\right)w = 0 \quad \text{and} \quad p_2\left(\frac{d}{dt}\right)w = 0, \quad p_1, p_2 \in \mathbb{R}[\xi].$$

Obtain a minimal kernel representation for $\mathcal{B}_1 \cap \mathcal{B}_2$ and $\mathcal{B}_1 + \mathcal{B}_2$ using the gcd and the lcm of p_1, p_2 .

3. Generalize to

$$p_1\left(\frac{d}{dt}\right)w = 0, \dots, p_n\left(\frac{d}{dt}\right)w = 0, \quad p_1, \dots, p_n \in \mathbb{R}[\xi].$$

4. Assume that \mathcal{B} is defined as the solution set of an *infinite* number of differential equations

$$R_\alpha\left(\frac{d}{dt}\right)w = 0, \quad \alpha \in \mathbb{A}, \quad R_\alpha \in \mathbb{R}[\xi]^{1 \times w},$$

with \mathbb{A} any (countably or uncountably) infinite set. Prove that there exists a polynomial matrix $R \in \mathbb{R}[\xi]^{\bullet \times w}$ (hence with a finite(!) number of rows) such that \mathcal{B} is specified by

$$R\left(\frac{d}{dt}\right)w = 0.$$

Exercise 12: Time-reversibility

Define the *time-reverse* of $\Sigma = (\mathbb{R}, \mathbb{W}, \mathcal{B})$ by

$$\text{reverse}(\Sigma) = (\mathbb{R}, \mathbb{W}, \text{reverse}(\mathcal{B}))$$

with

$$\text{reverse}(w)(t) := w(-t).$$

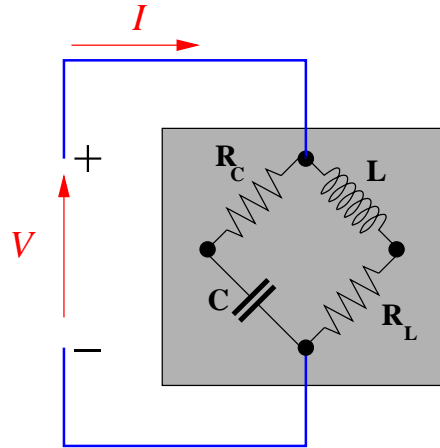
$\Sigma = (\mathbb{R}, \mathbb{W}, \mathcal{B})$ is said to be **time-reversible** if

$$\Sigma = \text{reverse}(\Sigma).$$

1. Prove that if $R\left(\frac{d}{dt}\right)w = 0$ is kernel representation of $\Sigma \in \mathcal{L}^w$, then $R\left(-\frac{d}{dt}\right)w = 0$ is a kernel representation of $\text{reverse}(\Sigma)$.
2. Do Kepler's laws define a time-reversible system?
3. Prove that $w + \frac{d^2}{dt^2}w = 0$ defines a time-reversible system.
4. Prove that the scalar system $p\left(\frac{d}{dt}\right)w = 0$ is time-reversible if and only if $p \in \mathbb{R}[\xi]$ is either an even or an odd polynomial.
5. Prove that the single-input/single-output system $p\left(\frac{d}{dt}\right)w_1 = q\left(\frac{d}{dt}\right)w_2$ is time-reversible if and only if $p, q \in \mathbb{R}[\xi]$ are both even or both odd.
6. Prove that the controllable single-input/single-output system $p\left(\frac{d}{dt}\right)w_1 = q\left(\frac{d}{dt}\right)w_2$ is time-reversible if and only if $p, q \in \mathbb{R}[\xi]$ are both even.

Exercise 13: Elimination of latent variables in the RLC circuit

Consider the RLC circuit discussed in the lecture on latent variables.



The equations that describe this circuit are KCL, the constitutive equations, and the manifest variable assignment.

1. Eliminate I_a, I_b, I_c, I_d and V_1, V_2, V_3, V_4 , and arrive at

$$C \frac{d}{dt} V = I_e + CR_C \frac{d}{dt} I_e, \quad V = R_L I_f + L \frac{d}{dt} I_f, \quad I = I_e + I_f.$$

Argue the correctness of these equations from first principles.

2. Next, eliminate I_f , and obtain

$$C \frac{d}{dt} V = I_e + CR_C \frac{d}{dt} I_e, \quad \frac{L}{R_L} \frac{d}{dt} I + I - \frac{1}{R_L} V = I_e + \frac{L}{R_L} \frac{d}{dt} I_e.$$

3. Finally, distinguish two cases, eliminate I_e , and derive the following differential equation governing (V, I) .

- ▶ For $CR_C \neq \frac{L}{R_L}$,

$$\begin{aligned} \left(\frac{R_C}{R_L} + \left(1 + \frac{R_C}{R_L} \right) CR_C \frac{d}{dt} + CR_C \frac{L}{R_L} \frac{d^2}{dt^2} \right) V \\ = \left(1 + CR_C \frac{d}{dt} \right) \left(1 + \frac{L}{R_L} \frac{d}{dt} \right) R_C I. \end{aligned}$$

- ▶ For $CR_C = \frac{L}{R_L}$, $\left(\frac{R_C}{R_L} + CR_C \frac{d}{dt} \right) V = \left(1 + CR_C \frac{d}{dt} \right) R_C I.$

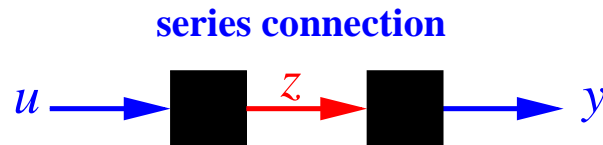
Exercise 14: Series and parallel connection

1. Use the Bézout identity to find a unimodular pre-multiplication that brings $\begin{bmatrix} p \\ q \end{bmatrix}$ with $p, q \in \mathbb{R}[\xi]$ into Smith form.
 p and q need not be coprime.



Étienne Bézout (1730-1783)

2. Consider the series connection of two single-input/single-output LTIDSs, as shown in the figure below.

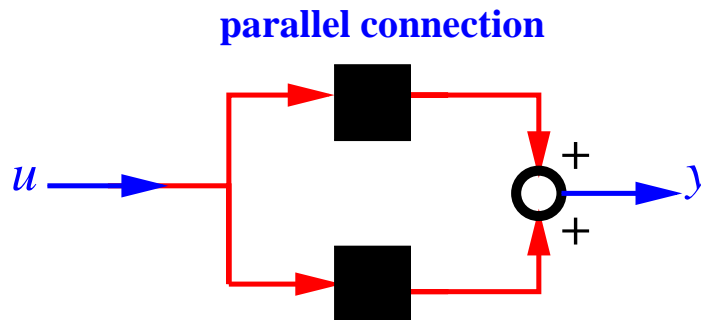


Assume that the systems are governed by respectively

$$p\left(\frac{d}{dt}\right)z = q\left(\frac{d}{dt}\right)u, \quad d\left(\frac{d}{dt}\right)y = n\left(\frac{d}{dt}\right)z.$$

Eliminate z to obtain a kernel representation for (u, y) .

3. Repeat 2. for parallel connection.



Exercise 15: Properties of controllable systems

1. Prove that

$$\begin{aligned} \llbracket \mathcal{B}_1, \mathcal{B}_2 \in \mathcal{L}^w \text{ controllable, } m(\mathcal{B}_1) = m(\mathcal{B}_2), \text{ and } \mathcal{B}_1 \subseteq \mathcal{B}_2 \rrbracket \\ \Rightarrow \llbracket \mathcal{B}_1 = \mathcal{B}_2 \rrbracket. \end{aligned}$$

Prove that the above implication does not hold without the controllability assumption.

2. Let $\mathcal{B}_1, \mathcal{B}_2 \in \mathcal{L}^2, m(\mathcal{B}_1) = m(\mathcal{B}_2) = 1, p(\mathcal{B}_1) = p(\mathcal{B}_2) = 1$
(hence both systems are single-input/single output systems).

Prove that

$$\llbracket \mathcal{B}_1, \mathcal{B}_2 \in \mathcal{L}^w \text{ controllable, } \mathcal{B}_1 \neq \mathcal{B}_2 \rrbracket \Rightarrow \llbracket \mathcal{B}_1 \cap \mathcal{B}_2 \text{ is autonomous} \rrbracket.$$

Prove that the above implication does not hold without the controllability assumption.

Exercise 16: Common factors

The coprimeness of polynomials has many applications in mathematics and has been studied since the time of Euler. Behavioral controllability proudly adds a new application of coprimeness.

Indeed, the LTIDS described by

$$a \left(\frac{d}{dt} \right) w_1 = b \left(\frac{d}{dt} \right) w_2,$$

with $a, b \in \mathbb{R}[\xi]$ is controllable (in the behavioral sense) if and only if a and b are coprime.

There are many algorithms for checking whether two polynomials $a, b \in \mathbb{R}[\xi]$ are coprime. Hundreds of articles has been published on this matter. These algorithms and their interrelations form the sort of mathematical topic that one can easily become addicted to. While it keeps mathematicians from doing mischief, it is advisable to show restraint in the pursuit of such mathematical puzzles. This exercise indulges a bit and discusses several algorithms for coprimeness.

Let $a, b \in \mathbb{R}[\xi]$. Written out in terms of their coefficients, we have

$$\begin{aligned} a(\xi) &= a_0 + a_1\xi + a_2\xi^2 + \cdots + a_{n-1}\xi^{n-1} + a_n\xi^n, \\ b(\xi) &= b_0 + b_1\xi + b_2\xi^2 + \cdots + b_{m-1}\xi^{m-1} + b_m\xi^m. \end{aligned}$$

In the present exercise, we obtain several test on the coefficients of a and b for checking coprimeness. Assume without loss of generality that $a_n \neq 0$ and $n \geq m$.



Leonhard Euler (1707-1783)

1. The Euclidean algorithm

Euclides already developed an algorithm for finding the gcd of two integers. This algorithm readily generalizes to polynomials.

Set $x_1 = a, x_2 = b$ and obtain x_k recursively by dividing with rest

$$x_k = x_{k+1}d + x_{k+2},$$

with $\text{degree}(x_{k+2}) < \text{degree}(x_{k+1})$ or $x_{k+2} = 0$.

Let $x_{k'}$ be the first time that $x_{k'} = 0$. **Prove that $x_{k'-1}$ is the gcd of a and b . Hence a and b are coprime if and only if $x_{k'} = 1$ for some k' .**

2. The Sylvester resultant

Let $\text{degree}(a) = n$ and $\text{degree}(b) = m$. Form the $(n+m) \times (n+m)$ matrix

$$\mathbb{S}(a, b) = \begin{matrix} \left[\begin{array}{cccccccc} a_0 & a_1 & \cdots & a_{n-1} & a_n & 0 & \cdots & \\ 0 & a_0 & \cdots & a_{n-2} & a_{n-1} & a_n & 0 & \cdots \\ & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots \\ & \cdots & 0 & a_0 & a_1 & a_2 & \cdots & a_n & 0 \\ & & \cdots & 0 & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_{m-1} & b_m & 0 & \cdots & & \\ 0 & b_0 & \cdots & b_{m-2} & b_{m-1} & b_m & 0 & \cdots & \\ & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \\ & \cdots & 0 & b_0 & b_1 & b_2 & \cdots & b_m & 0 \\ & & \cdots & 0 & b_0 & b_1 & b_2 & \cdots & b_m \end{array} \right] \end{matrix} \left. \begin{array}{l} \vphantom{\left[\right.} \right\} \text{m rows} \\ \vphantom{\left[\right.} \right\} \text{n rows} \end{array} \right.$$

$\mathbb{S}(a, b)$ is called the *Sylvester matrix*, and its determinant is called the *Sylvester resultant* of a and b .

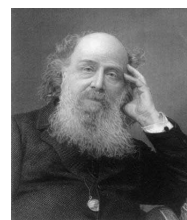
Prove that the polynomials a and b are coprime if and only if the Sylvester resultant of a and b is non-zero.

Hint: Consider the map $(x, y) \mapsto ax + by$.

View this as a map from the real polynomials with $\text{degree}(x) < m$ and $\text{degree}(y) < n$ to the real polynomials with $\text{degree} < n+m$. The map $(x, y) \mapsto$



Euclides (325-265 BC)
painting by Raffaello



James Sylvester (1814-1897)

$ax + by$ thus becomes a map from \mathbb{R}^{n+m} to \mathbb{R}^{n+m} . Prove that it is injective if and only a and b are coprime. Now write this map in matrix notation and obtain the Sylvester test for coprimeness.

3. The Hankel matrix

Another test for coprimeness uses the ‘Laurent series’ expansion of $g = \frac{b}{a} \in \mathbb{R}(\xi)$ defined by

$$g(\xi) = \frac{b(\xi)}{a(\xi)} = g_0 + \frac{g_1}{\xi} + \frac{g_2}{\xi^2} + \cdots + \frac{g_n}{\xi^n} + \cdots.$$

Form the ‘Hankel matrix’

$$\mathbb{H}(g)_{n \times n} = \begin{bmatrix} g_1 & g_2 & \cdots & g_{n-1} & g_n \\ g_2 & g_3 & \cdots & g_n & g_{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{n-1} & g_n & \cdots & g_{2n-3} & g_{2n-2} \\ g_n & g_{n+1} & \cdots & g_{2n-2} & g_{2n-1} \end{bmatrix}.$$



Hermann Hankel (1839-1873)

Prove that a and b are coprime if and only if $\det(\mathbb{H}(g)_{n \times n}) \neq 0$.

Hint: From the preamble, we can conclude that a and b are coprime if and only if the first n columns of $\mathbb{H}(g)$ are linearly independent. Furthermore the $(n+k)$ -th column of $\mathbb{H}(g)$ is linearly dependent on the first n columns for $k = 1, 2, \dots$. Therefore, by the Hankel structure, the $(n+k)$ -th rows of the matrix formed by the first n columns of $\mathbb{H}(g)$ are linearly dependent on the first n rows for $k = 1, 2, \dots$. Conclude that a and b are coprime if and only if the first n rows of the matrix formed by the first n columns of $\mathbb{H}(g)$ has rank n .

4. The Bézoutian

Form the bivariate polynomial

$$B(a,b)(\zeta, \eta) = \frac{a(\zeta)b(\eta) - a(\eta)b(\zeta)}{\zeta - \eta}.$$

$B(a,b)$ is called the *Bézoutian* of the polynomials a and b .

Prove that $B(a,b) \in \mathbb{R}[\zeta, \eta]$.

Hint: $\frac{M(\zeta, \eta)}{\zeta - \eta}$ with $M \in \mathbb{R}[\zeta, \eta]$ is polynomial if and only if $M(\xi, \xi) = 0$.

$B(a, b)(\zeta, \eta)$ can be written in matrix notation as

$$B(a, b)(\zeta, \eta) = \begin{bmatrix} 1 \\ \zeta \\ \zeta^2 \\ \vdots \\ \zeta^{n-1} \end{bmatrix} \mathbb{B}(a, b) \begin{bmatrix} 1 \\ \eta \\ \eta^2 \\ \vdots \\ \eta^{n-1} \end{bmatrix}.$$



Étienne Bézout (1730-1783)

The determinant of $\mathbb{B}(a, b)$ is called the *Bézout resultant* of a and b .

Prove that $\mathbb{B}(a, b) = \mathbb{B}(a, b)^\top \in \mathbb{R}^{n \times n}$ and that the Bézout resultant of a and b is non-zero if and only if a and b are coprime.

Hint: The following relation holds between the Hankel matrix and the Bézoutian

$$\mathbb{B}(a, b) = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ a_2 & a_3 & \cdots & a_n & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1} & a_n & \cdots & 0 & 0 \\ a_n & 0 & \cdots & 0 & 0 \end{bmatrix} \mathbb{H}(g)_{n \times n} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ a_2 & a_3 & \cdots & a_n & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1} & a_n & \cdots & 0 & 0 \\ a_n & 0 & \cdots & 0 & 0 \end{bmatrix}$$

In order to prove this, proceed as follows.

(i) Prove that

$$B(\zeta, \eta) = a(\zeta) \frac{g(\eta) - g(\zeta)}{\zeta - \eta} a(\eta).$$

(ii) From there, arrive at

$$B(\zeta, \eta) = \sum_{k, \ell=1, 2, \dots, n} g_{k+\ell-1} (a_n \zeta^{n-k} + \cdots + a_0 \zeta^{-k}) (a_n \eta^{n-\ell} + \cdots + a_0 \eta^{-\ell}).$$

(iii) Conclude that

$$B(\zeta, \eta) = \sum_{k, \ell=1, 2, \dots, n} g_{k+\ell-1} (a_n \zeta^{n-k} + \cdots + a_k) (a_n \eta^{n-\ell} + \cdots + a_\ell).$$

5. The MacDuffee resultant

As if four tests for coprimeness do not yet suffice, we now give a fifth one, due to MacDuffee. Define the companion matrix of a by

$$A_{\text{companion}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \cdots & -\frac{a_{n-2}}{a_n} & -\frac{a_{n-1}}{a_n} \end{bmatrix}$$

Prove that a and b are coprime if and only if $\det(b(A_{\text{companion}}))$, the MacDuffee resultant of a and b , is not zero.

Hint: Assume first that the roots of a are distinct, and obtain the MacDuffee test by diagonalizing $A_{\text{companion}}$. Then obtain the general case by a continuity argument.

State and prove an analogous test with the roles of a and b reversed.

Exercise 17: Observable image representations

The image representation $w = M\left(\frac{d}{dt}\right)\ell$ is said to be **observable** if

$$\llbracket M\left(\frac{d}{dt}\right)\ell_1 = M\left(\frac{d}{dt}\right)\ell_2 \rrbracket \Leftrightarrow \llbracket \ell_1 = \ell_2 \rrbracket.$$

1. Prove that a controllable system $\mathcal{B} \in \mathcal{L}^w$ admits an observable image representation.
2. Assume that the single-input/single-output system $p\left(\frac{d}{dt}\right)w_1 = q\left(\frac{d}{dt}\right)w_2$ is controllable. Give an observable image representation for this system.

Exercise 18: Non-anticipation

1. Consider the input/output system $\Sigma = (\mathbb{R}, \mathbb{R}^m \times \mathbb{R}^p, \mathcal{B})$ defined by

$$y = G \left(\frac{d}{dt} \right) u \quad \text{with } G \in \mathbb{R}(\xi)^{p \times m}.$$

We say that y **does not anticipate** u if for all $(u, y) \in \mathcal{B}$ and $u' \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^m)$ such that $u'(t) = u(t)$ for $t \leq 0$, there exists $y' \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^p)$ such that $(u', y') \in \mathcal{B}$ and $y'(t) = y(t)$ for $t \leq 0$.

Prove that y does not anticipate u (without conditions on G !).

2. Consider now the discrete-time input/output system $\Sigma = (\mathbb{Z}, \mathbb{R}^m \times \mathbb{R}^p, \mathcal{B})$ defined by

$$y = G(\sigma)u \quad \text{with } G \in \mathbb{R}(\xi)^{p \times m}.$$

Define non-anticipation. Prove that y does not anticipate u if and only if G is proper.

3. Does the moving average system

$$y(t) = \frac{1}{2N+1} \sum_{t'=-N}^N u(t+t')$$

define a non-anticipating system?

4. For what $\Delta \in \mathbb{R}$ does the differential-delay system

$$\frac{d}{dt}y(t) = u(t + \Delta)$$

define a nonanticipating system?

Exercise 19: Norm-preserving representations

The representation of this exercise uses the following ‘spectral factorization’-like result.

Assume that $P \in \mathbb{R}[\xi]^{n \times n}$ satisfies $P(\xi) = P(-\xi)^\top$ and $P(i\omega) > 0$ for $\omega \in \mathbb{R}$. Then there exists $F \in \mathbb{R}[\xi]^{n \times n}$ such that $P(\xi) = F(-\xi)^\top F(\xi)$.

1. Prove this factorization for the case $n = 1$.
2. Prove that if $\mathcal{B} \in \mathcal{L}^w$ is controllable, then it admits an observable ‘image’ representation

$$w = N \left(\frac{d}{dt} \right) \ell$$

with $N \in \mathbb{R}(\xi)^{w(\mathcal{B}) \times m(\mathcal{B})}$ such that

$$N(-\xi)N(\xi) = I.$$

Proceed as follows. Start with an observable image representation $w = M \left(\frac{d}{dt} \right) \ell$ with $M \in \mathbb{R}[\xi]^{w(\mathcal{B}) \times m(\mathcal{B})}$. Then factor $M^\top(-\xi)M(\xi)$ as $F^\top(-\xi)F(\xi)$ with $F \in \mathbb{R}[\xi]^{m(\mathcal{B}) \times m(\mathcal{B})}$. Take $N = MF^{-1}$.

3. Prove that this representation has the property that

$$\|w\|_{\mathcal{L}_2(\mathbb{R}, \mathbb{R}^w)} = \|\ell\|_{\mathcal{L}_2(\mathbb{R}, \mathbb{R}^m)},$$

hence the name ‘norm-preserving’ image representation.

4. Prove that such a representation exists only very exceptionally with N polynomial. Norm-preserving representations have a number of applications, and require rational symbols.

Exercise 20: The structure of \mathcal{L}^\bullet

Let $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2 \in \mathcal{L}^\bullet$ and $F \in \mathbb{R}[\xi]^{\bullet \times \bullet}$.

1. Prove that $(\mathcal{B}_1 + \mathcal{B}_2) \in \mathcal{L}^\bullet$.

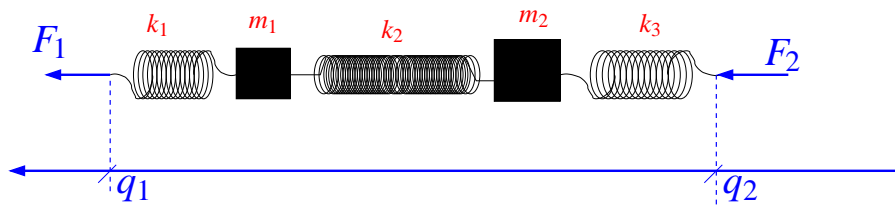
2. Prove that $\mathcal{B}_1 \cap \mathcal{B}_2 \in \mathcal{L}^\bullet$.

3. Prove that $F \left(\frac{d}{dt} \right) \mathcal{B} \in \mathcal{L}^\bullet$.

4. Prove that $F \left(\frac{d}{dt} \right)^{-1} \mathcal{B} \in \mathcal{L}^\bullet$.

Assume, as always with the \bullet notation, that the relevant objects have compatible dimensions.

Exercise 21: Modeling a mass-spring-damper system

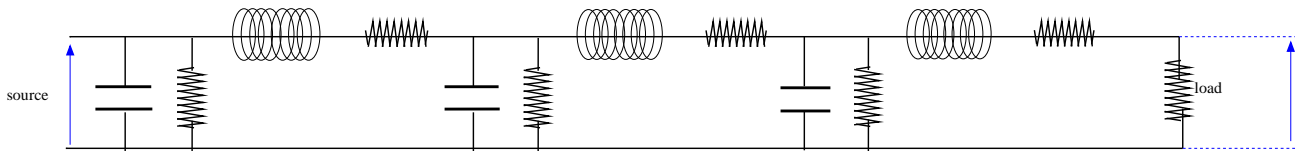


Consider the mass-spring device shown above. Assume that it operates horizontally from equilibrium in its linear mode. The problem is to model the relation between the forces F_1, F_2 , and the horizontal positions q_1, q_2 .

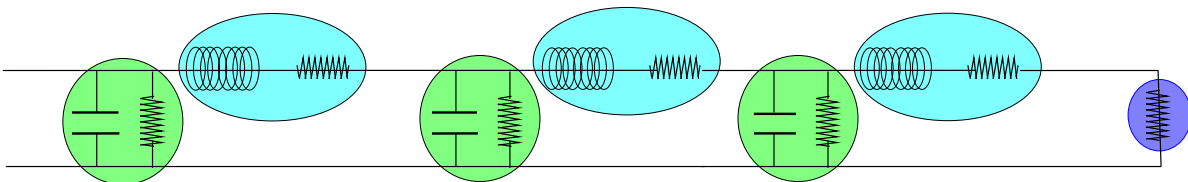
1. View this system as consisting of 5 subsystems. Define the graph with leaves that determines the interconnection architecture.
2. Choose latent and manifest variables, and write the module equations.
3. Write the interconnection laws.
4. Eliminate the latent variables and obtain behavioral equations involving only the manifest variables.

Exercise 22: Modeling a transmission line

Consider the transmission line modeling problem discussed during Lecture VIII.



1. View the transmission line as an interconnection of 7 subsystems as shown below.



Determine the graph with leaves that defines the interconnection architecture.

2. There are 3 kinds of subsystems: **blue**, **green**, and **cyan**. Model each of these subsystems.
3. Specify the interconnection laws.
4. Specify the manifest variables.
5. Obtain the full set of equations leading, after elimination of the latent variables, to the desired differential equation that describes the behavior of (w_1, w_2)

$$r_1 \left(\frac{d}{dt} \right) w_1 = r_2 \left(\frac{d}{dt} \right) w_2.$$

Exercise 23: Characteristic polynomial assignment

Consider the plant $\mathcal{P} \in \mathcal{L}^w$. As all LTIDSs, it can be decomposed into

$$\mathcal{P} = \mathcal{P}_{\text{controllable}} \oplus \mathcal{P}_{\text{autonomous}}.$$

Let $\mathcal{C} \in \mathcal{L}^w$ be a regular controller and consider the autonomous controlled system $\mathcal{P} \cap \mathcal{C} \in \mathcal{L}^w$.

1. Prove that for any monic $\pi \in \mathbb{R}[\xi]$ there exists such a \mathcal{C} such that

$$\chi_{\mathcal{P} \cap \mathcal{C}} = \pi$$

if and only if $\chi_{\mathcal{P}_{\text{autonomous}}}$ is a factor of π . Deduce the pole placement theorem from here.

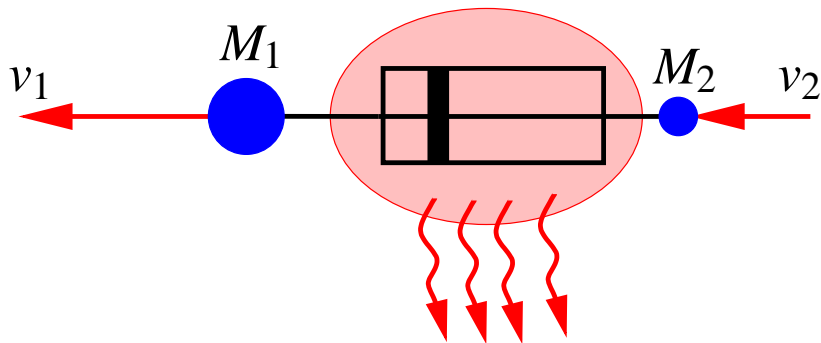
2. Repeat the same question with the characteristic polynomial replaced by the minimal polynomial.

Exercise 24: Preservation of properties under interconnection

The notation is the one used in Lectures IV and V. \mathcal{B} is the original behavior, \mathcal{B}' the behavior obtained after interconnection of terminals.

1. Consider the interconnection of terminals of an electrical circuit. Prove that if \mathcal{B} is linear and time-invariant, so is \mathcal{B}' . Prove that if \mathcal{B} is a LTIDS, so is \mathcal{B}' . Prove that if \mathcal{B} satisfies KVL, so does \mathcal{B}' . Prove that if \mathcal{B} satisfies KCL, so does \mathcal{B}' .
2. Consider the interconnection of terminals of a mechanical system. Prove that if \mathcal{B} satisfies IUM, so does \mathcal{B}' .
3. Consider Newton's second law. Prove IUM.
4. Apparently, Kepler's laws do not satisfy IUM. Discuss why this is and in what sense IUM holds for Kepler's laws.

Exercise 25: Heat produced by a damper



Consider the system shown above. It consists of 2 masses, connected by a damper. The motion is assumed to take place horizontally. The damper is assumed to be a linear damper with damping coefficient $D > 0$.

1. Denote the positions of the pointmasses by q_1, q_2 . There are no external forces that act on the masses. Obtain the differential equations that govern (q_1, q_2) .
2. Solve this differential equation with initial conditions $q_1(0), q_2(0), \frac{d}{dt}q_1(0), \frac{d}{dt}q_2(0)$
3. What is the power going into the damper?
4. Assume that all the energy absorbed by the damper is converted into heat. Compute the energy absorbed on the interval $[0, \infty)$ as a function of the parameters $M_1, M_2, D, q_1(0), q_2(0), \frac{d}{dt}q_1(0), \frac{d}{dt}q_2(0)$. Note that this energy only depends on $M_1, M_2, \frac{d}{dt}q_1(0), \frac{d}{dt}q_2(0)$.