Solutions Exercises Set 3

Exercise 1 (Controllability and interconnection)

The aim of this exercise is (i) to illustrate the behavioral concept of controllability and (ii) to show its fragility under system operations.

1. Let $\mathfrak{B}_1 \in \mathfrak{L}^\bullet$ be described by

$$R_1(\frac{d}{dt})w_1 = R_2(\frac{d}{dt})w_2, \tag{1}$$

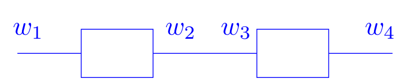
and $\mathfrak{B}_2 \in \mathfrak{L}^\bullet$ be described by

$$R_3(\frac{d}{dt})w_3 = R_4(\frac{d}{dt})w_4. \tag{2}$$

Define their *series (or cascade) interconnection* by the full behavioral equations (1,2) combined with

$$w_2 = w_3.$$

Of course, we assume that the dimensions are such that this equation makes sense. In the manifest behavior, consider (w_1, w_4) as the manifest variables and (w_2, w_3) as latent variables (i.e., in behavioral equations for the manifest behavior of the series connection, w_2 and w_3 are eliminated).



Consider the system with transfer function $\frac{1}{s}$, i.e.,

$$\frac{d}{dt}y_1 = u_1$$

and the system with transfer function s, i.e.,

$$y_2 = \frac{d}{dt}u_2.$$

Are these systems controllable? Compute behavioral equations for the manifest behavior of the series connection defined by $u_2 = y_1$. Is this system controllable? What is its transfer function? Now consider the series connection in opposite order, i.e., defined by $u_1 = y_2$. Compute behavioral equations for the manifest behavior of this series connection. Is this system controllable? What is its transfer function? Are the two series connections the same? If not, give a signal that belongs to the manifest behavior of one, but not the other. Does series connection of single-input/single-output connections 'commute'?

(1) $\Leftrightarrow R(\frac{d}{dt})\begin{bmatrix}u_1\\y_1\end{bmatrix} = 0$ **with** $R(\xi) = [-1 \quad \xi]$; $\operatorname{rank}(R(\lambda)) = 1$ **for all** $\lambda \in \mathbb{C} \Rightarrow$ **controllable.**

(2) $\Leftrightarrow R(\frac{d}{dt})\begin{bmatrix}u_2\\y_2\end{bmatrix} = 0$ **with** $R(\xi) = [-\xi \quad 1]$; $\operatorname{rank}(R(\lambda)) = 1$ **for all** $\lambda \in \mathbb{C} \Rightarrow$ **controllable.**

series connection: $\quad -\boxed{\tfrac{1}{s}}-\boxed{s}- \quad \equiv \quad -\boxed{1}-$

$$\frac{d}{dt}y_1 = u_1, \qquad y_2 = \frac{d}{dt}u_2, \qquad u_2 = y_1$$

Eliminating y_1 **and** u_2**, yields the system** $u_1 = y_2$**, which is obviously controllable. The transfer function is 1.**

Consider now the series connection in the opposite order: $\quad -\boxed{s}-\boxed{\tfrac{1}{s}}- \quad \equiv \quad -\boxed{\tfrac{s}{s}}-$

$$\frac{d}{dt}y_1 = u_1, \qquad y_2 = \frac{d}{dt}u_2, \qquad u_1 = y_2$$

Eliminating y_2 **and** u_1**, yields the system**

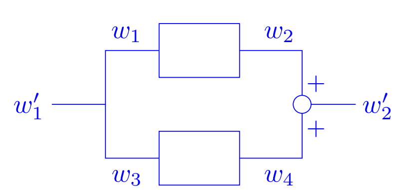
$$\frac{d}{dt}y_1 = \frac{d}{dt}u_2 \quad \Leftrightarrow \quad R(\frac{d}{dt})\begin{bmatrix}u_2\\y_1\end{bmatrix} = 0, \qquad \textbf{with } R(\xi) = [-\xi \quad \xi]$$

$\operatorname{rank}(R(\lambda)) = 1$ **for** $0 \neq \lambda \in \mathbb{C}$**, and** $\operatorname{rank}(R(\lambda)) = 0$ **for** $\lambda = 0$**, so the system is not controllable. The transfer function is** $\xi^{-1}\xi = 1$**.**

The two series connections are not equivalent, even though they have the same transfer function. Any non-zero constant input-output belongs to the second series connection, but not to the first. Hence series connection does not commute. It does commute, though, for the transfer functions, i.e., for the controllable part.

Comment: When we write $\frac{s}{s}$ for a transfer function, or, generally, a transfer function with a common factor in the numerator and denominator, we mean exactly the same as $\frac{1}{1}$, with the common factor cancelled. Indeed, in rational functions one can *by definition of a rational function* cancel (or add) common factors. So, when you read or hear: *assume that there are no common factors in the numerator and denominator of this or that transfer function*, smile, and think *'innocence is bliss'*. What this assumption usually means is that people actually have a kernel representation, in which lack of common factors means controllability. But since they have been brought up without the notion of kernel representation, but with the thought that a system IS a transfer function, they have to resort to such convoluted statements involving common factors.

2. Define, in the above spirit of series connection, *parallel connection*.



Parallel connection: (1,2) combined with

$$w_1' = w_1 = w_3, \qquad w_2' = w_2 + w_4,$$

with w_1' **and** w_2' **the manifest variables and** w_1**,** w_2**,** w_3**, and** w_4**, the latent variables.**

3. Decide, by means of a proof or a counterexample, which of the above operations preserve controllability. Of course, we assume that we deal with systems in $\mathfrak{L}^\bullet$, and that the dimensions are appropriate:

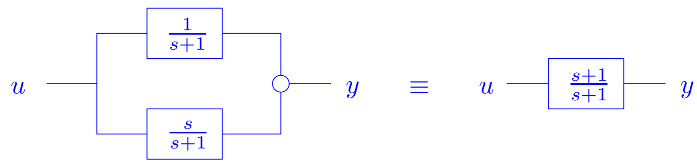
   (a) series connection

   Series connection does not preserve controllability, see 1.

   (b) parallel connection

   Parallel connection connection does not preserve controllability.

   Example:



$$\left(\frac{d}{dt}+1\right)y_1 = u_1, \qquad \left(\frac{d}{dt}+1\right)y_2 = \frac{d}{dt}u_2, \qquad u = u_1 = u_2, \qquad y = y_1 + y_2$$

   After elimination: $(\frac{d}{dt}+1)y = (\frac{d}{dt}+1)u$, **so** $R(\frac{d}{dt})\begin{bmatrix} u \\ y \end{bmatrix} = 0$, **with** $R(\xi) = [-\xi - 1 \quad \xi + 1]$, **which drops rank for** $\lambda = -1$.

   (c) addition, i.e., does $\mathfrak{B}_1, \mathfrak{B}_2$ controllable imply $\mathfrak{B}_1 + \mathfrak{B}_2$ controllable?

   Define $\mathfrak{B}_1 + \mathfrak{B}_2$ **by,**

$$\mathfrak{B}_1: \ R_1(\frac{d}{dt})w_1 = 0, \qquad \mathfrak{B}_2: \ R_2(\frac{d}{dt})w_2 = 0, \qquad \mathfrak{B}_1 + \mathfrak{B}_2: \ w = w_1 + w_2,$$

   with w_1, w_2 **latent variables, and** w **the manifest one.** $\mathfrak{B}_1$ **and** $\mathfrak{B}_2$ **are controllable iff the full behavior is controllable, which implies that** $\mathfrak{B}_1 + \mathfrak{B}_2$ **is controllable (elimination preserves controllability, see 3e).**

   (d) intersection

   Let

$$\mathfrak{B}_1: \ R_1(\frac{d}{dt})w_1 = 0, \qquad \mathfrak{B}_2: \ R_2(\frac{d}{dt})w_2 = 0.$$

   The intersection of $\mathfrak{B}_1$ **and** $\mathfrak{B}_2$ **does not preserve controllability.**

   Take, for example, $R_1 = [p_1 \quad q_1]$, $R_2 = [p_2 \quad q_2]$, $\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \end{bmatrix}$ **drops rank at the roots of** $p_1 q_2 - q_1 p_2$.

   (e) elimination

   Elimination preserves controllability. Go back to the basic definition of controllability for a straightforward proof, that is also valid for nonlinear systems. Or consider the representation $R(\frac{d}{dt})w = Monl$, **use a unimodular pre-multiplication, if necessary, to write this as** $R'(\frac{d}{dt})w = 0, R"(\frac{d}{dt})w = M"(\frac{d}{dt})\ell$, **with** $M"$ **of full row rank. Note that** $R'(\frac{d}{dt})w = 0$ **is a kernel representation of the manifest behavior. Finally, observe that rank constancy of** $[R(\lambda) \quad M(\lambda)]$ **implies rank constancy of** $R'(\lambda)$. **Hence controllability of the full behavior implies controllability of the manifest behavior.**

   (f) action of a linear differential operator i.e., does $\mathfrak{B}$ controllable and $F \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ imply $F(\frac{d}{dt})\mathfrak{B}$ controllable?

   Assume that $\mathfrak{B}$ **is controllable. Hence it admits an image representation** $w = M(\frac{d}{dt})\ell$.
   It follows that $F(\frac{d}{dt})\mathfrak{B}$ **(variables** v**) is defined by** $v = F(\frac{d}{dt})w = F(\frac{d}{dt})M(\frac{d}{dt})\ell$.
   Hence $F(\frac{d}{dt})\mathfrak{B}$ **admits an image representation, so that** $F(\frac{d}{dt})\mathfrak{B}$ **is controllable.**

(g) the inverse action of a linear differential i.e., does $\mathfrak{B}$ controllable and $F \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ imply that $\{w \mid F(\frac{d}{dt})w \in \mathfrak{B}\}$ (this is a system in $\mathfrak{L}^{\bullet}$!) is controllable?

Assume that $\mathfrak{B}$ is controllable and has a representation $R(\frac{d}{dt})v = 0$. Then $F^{-1}(\frac{d}{dt})\mathfrak{B}$ is governed by

$$F(\frac{d}{dt})w \in \mathfrak{B}, \text{ i.e., } R(\frac{d}{dt})F(\frac{d}{dt})w = 0.$$

Clearly $R(\lambda)$ could have constant rank, but $R(\lambda)F(\lambda)$ not (take for example $R(\xi) = \xi$). Hence $F^{-1}(\frac{d}{dt})\mathfrak{B}$ need not be controllable.

Exercise 2 (Moving average)

The aim of this exercise is to illustrate that the notion of controllability can even shed some light on some very common algorithms.

Throughout this exercise, the time-axis is $\mathbb{Z}$.

1. (Recall first exercise 2 of set 2). Let σ denote, as usual, the *shift*: $\sigma(f)(t) := f(t+1)$. Let $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi, \xi^{-1}]$ and consider the system of difference equations
$$R(\sigma, \sigma^{-1})w = 0.$$
This defines the dynamical system $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$. Prove that this system is controllable iff
$$\operatorname{rank}(R(\lambda, \lambda^{-1})) = \text{ constant over } 0 \neq \lambda \in \mathbb{C}.$$
Prove by an example that you cannot dispense of 'puncturing' 0 from $\mathbb{C}$ in this test.

The Smith form for matrices over $\mathbb{R}[\xi, \xi^{-1}]$ reads: Let $M \in \mathbb{R}^{\bullet \times \bullet}[\xi, \xi^{-1}]$. There exist unimodular $U, V \in \mathbb{R}^{\bullet \times \bullet}[\xi, \xi^{-1}]$, such that UMV is of the form
$$UMV = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$
with $D = \operatorname{diag}(d_1, d_2, \ldots, d_{\mathtt{r}})$, $d_{\mathtt{k}} \in \mathbb{R}[\xi, \xi^{-1}]$, and $d_{\mathtt{k}+1}$ is a factor of $d_{\mathtt{k}}$, for $\mathtt{k} = 1, \ldots, \mathtt{r}-1$. In fact, we can take $d_{\mathtt{k}} \in \mathbb{R}[\xi]$ with $d_{\mathtt{k}}(0) \neq 0$.

Now, proceed exactly as in the continuous-time case: $R(\sigma, \sigma^{-1})w = 0$ defines a controllable system iff $D(\sigma, \sigma^{-1})w = 0$ does. The latter is the case iff each of the systems $d_{\mathtt{k}}(\sigma, \sigma^{-1})w_{\mathtt{k}} = 0$ defines a controllable system. This is the case iff each of the $d_{\mathtt{k}}$'s is unimodular. Expressed in terms of R, this yields the rank condition.

Note finally that the puncturing is indeed necessary. Consider the system described by $\sigma w = 0$. i.e., $\mathfrak{B} = 0$. It is obviously controllable. The associated $R(\xi)$ is ξ. $R(\lambda)$ *drops rank at* $\lambda = 0$, but this does not contradict controllability.

2. Consider the system defined by
$$w_2(t) = \frac{1}{T}\, \Sigma_{t'=1,2,\ldots,T}\, w_1(t - t'). \tag{3}$$
This algorithm is called a *moving average (MA)* smoothing. $T \in \mathbb{N}$ is called the *averaging window*. It is very frequently used in order to filter out noise, detecting trends, etc. When T is large, it is tempting to replace this algorithm by
$$w_2(t) = w_2(t-1) + \frac{1}{T}(w_1(t-1) - w_1(t-T-1)). \tag{4}$$

   (a) Do (3) and (4) have the same transfer function?

   The transfer function $w_1 \mapsto w_2$ of (3) is
   $$G(\xi) = \frac{1}{T}(\xi^{-1} + \xi^{-2} + \cdots + \xi^{-T}) = \frac{1}{T}\frac{\xi^{-1} - \xi^{-(T+1)}}{1 - \xi^{-1}},$$
   and the transfer function of (4) is
   $$G(\xi) = \frac{1}{T}\frac{\xi^{-1} - \xi^{-(T+1)}}{1 - \xi^{-1}}.$$
   (3) and (4) have the same transfer function.

(b) Compare, by counting the number of additions and multiplications required per time-step, (3) and (4) from the computational complexity point of view.

Per "time step" (3) takes $T-1$ additions and 1 multiplication, while (4) takes only 2 additions and 1 multiplication. From this point of view, (4) seems simpler.

(c) Do (3) and (4) define the same system (of course, in the behavioral sense, the one and only way...)?

No, (3) is controllable and (4) is not (see 2d). For example, $w_1(t) = c_1$, $w_2(t) = c_2$ is a solution of (4), but not of (3) if $c_1 \neq c_2$.

(d) Is (3) controllable? Is (4) controllable?

Apply part 1: (3) is controllable since

$$R(\xi) = [-\frac{1}{T}(\xi^{-1} + \xi^{-2} + \cdots + \xi^{-T}), 1], \qquad \operatorname{rank}(R(\lambda)) = 1, \quad \forall\ 0 \neq \lambda \in \mathbb{C}.$$

For (4), we get

$$R(\xi) = [-\frac{1}{T}(1 - \xi^{-(T+1)}), 1 - \xi^{-1}], \qquad \operatorname{rank}(R(1)) = 0,$$

so it is not controllable.

(e) Find a controllable $\oplus$ autonomous decomposition of (4).

$\mathfrak{B}_{\textbf{controllable}}$: **kernel representation (3)**
$\mathfrak{B}_{\textbf{autonomous}}$: **kernel representation** $w_1 = 0,\ \sigma w_2 = w_2$.

(f) Would you call (3) stable? (We use 'stable' as meaning $[w_1(t) = 0, (w_1, w_2) \in \mathfrak{B}] \Rightarrow [w_2(t) \to t \to \infty]$.) (4)? Does this conclusion make (4) useless as an algorithm?

(3) is stable: if $w_1 = 0$ for $t \geq 0$, then $w_2(t) \xrightarrow{t\to\infty} 0$. (4) is not stable: if $w_1 = 0$ for $t \geq 0$, then $w_2(t)$ does not necessarily go to zero (it may be a non-zero constant, see 2e). (4) is rather useless: if an error occurs in the calculations of $w_2(t')$, this error will appear in the results forever after that, for $t > t'$.

Conclude that stability is definitely not solely an issue about transfer functions, notwithstanding what is written in an endless number of papers in the IEEE Transactions on Automatic Control.

3. A very close relative of (3) is

$$w_2(t) = \frac{1}{2T+1} \sum_{t'=-T,\ldots,-1,0,1,\ldots,T} w_1(t - t'). \tag{5}$$

We have seen during the lectures that every system in $\mathfrak{L}^\bullet$ admits an componentwise input/output partition with a proper transfer function. This result generalizes to the discrete time case (you may want to prove this): every system admits a componentwise i/o partition in which the input does not anticipate the output (sometimes this is called *'causality'*). Consider the system defined by

$$p(\sigma, \sigma^{-1})y = q(\sigma, \sigma^{-1})u,$$

with $0 \neq p, q \in \mathbb{R}[\xi, \xi^{-1}]$. Determine in terms of the transfer function when the input does not anticipate the output. What is this non-anticipating input/output partition for (3)? Do you have a choice? Same question for (5). In the first case, the answer is what every sensible person would expect. In the second case the answer is a bit absurd. What is the explanation of this annoying aspect of (5)? The problem, it appears, is with the notions of input and output. They are much more context dependent that we have been led to believe.

First, we analyze when in the system with behavior $\mathfrak{B}$ described by

$$p(\sigma,\sigma^{-1})w_1 = q(\sigma,\sigma^{-1})w_2, \qquad 0 \neq p,q \in \mathbb{R}[\xi,\xi^{-1}],$$

w_1 does not anticipate w_2. Noting the results of exercise set 2, exercise 2, we may as well assume that p and q look like

$$p(\xi,\xi^{-1}) = 1 + p_1\xi^{-1} + \cdots, \qquad q(\xi,\xi^{-1}) = q_L\xi^L + q_{L-1}\xi^{L-1} + \cdots, \qquad q_L \neq 0.$$

The resulting difference equation looks like

$$w_1(t) = -p_1w_1(t-1) - p_2w_1(t-2) - \cdots + q_Lw_2(t+L) + q_{L-1}w_2(t+L-1) + \cdots$$

Now w_1 does not anticipate w_2 iff $w_2(t) = 0$ for $t < 0$ implies that there exists w_1, such that $(w_1,w_2) \in \mathfrak{B}$ and $w_1(t) = 0$ for $t < 0$. Note that w_2 is free in $\mathfrak{B}$. The above difference equation shows that w_1 does not anticipate w_2 iff $L \leq 0$, equivalently if the transfer function

$$G(\xi) = \frac{q(\xi,\xi^{-1})}{p(\xi,\xi^{-1})}$$

is proper (meaning $\lim_{\lambda\in\mathbb{R},\lambda\to\infty} |G(\lambda)| < \infty$).
Now consider (3). The transfer function $w_1 \mapsto w_2$ is

$$\frac{1}{T}\frac{\xi^{-1} - \xi^{-(T+1)}}{1-\xi^{-1}}.$$

Since this is a proper rational function w_1 does not anticipate w_2. (4) puts this into evidence. The transfer function $w_2 \mapsto w_1$ is

$$T\frac{\xi - 1}{1-\xi^{-T}},$$

which is not proper, so w_2 anticipates w_1. Thus, if we look for a causal input/output partition, we have no choice: w_1 is input and w_2 is output. This is all very reasonable.
Now consider (5). The transfer function $w_1 \mapsto w_2$ is

$$\frac{1}{2T+1}(\xi^T + \cdots + \xi + 1 + \xi^{-1} + \cdots + \xi^{-T}) = \frac{1}{2T+1}\frac{\xi^{T+1} - \xi^{-T}}{\xi - 1},$$

which is not a proper rational function, so w_1 anticipates w_2. The transfer function $w_2 \mapsto w_1$, on the other hand, is

$$(2T+1)\frac{\xi - 1}{\xi^{T+1} - \xi^{-T}},$$

which is a proper rational function, so w_2 does not anticipates w_1. Thus is we look for a causal input/output trajectories, we have again no choice: w_2 is input and w_1 is output. This is an absurd, of course, since, as a smoother, w_1 is the input in (5).

Comment: Of course, this "problem" occurs because (5) is a non-real-time smoother, and so we should not apply our control based system theoretic intuition in these situations. But such non-real-time algorithms are very much part of the trade in signal processing.

4. Another close relative of (3) is *exponential weighting:*

$$w_2(t) = \frac{1}{\rho^{-1}-1} \sum_{t' \in \mathbb{N}} \rho^{t'} w_1(t-t') \tag{6}$$

with $\rho \in (0,1)$ the *weighting parameter.* Convolutions as (6) or the continuous time analogs are of course very much related to our linear difference or differential systems. For the case at hand the associated difference equation is

$$w_2(t) = \rho w_2(t-1) + \frac{\rho}{\rho^{-1}-1}\, w_1(t-1) \tag{7}$$

(a) (6) has the drawback that it hard to give a very concrete characterization of the behavior, since it unclear for which $w_1 : \mathbb{Z} \to \mathbb{R}$ the infinite sum is finite. Prove that the infinite sum is finite when w_1 is bounded. Prove that (6) combined with w_1 bounded, and (7) combined with w_1, w_2 bounded have the same behavior.

Assume that (w_1, w_2) satisfies (6) and that w_1 is bounded. Denote by $\| \ \|_\infty$ the ℓ_∞-norm. Then

$$\begin{aligned}
|w_2(t)| &\leq \frac{1}{\rho^{-1}-1} \left| \sum_{t' \in \mathbb{N}} \rho^{t'} w_1(t-t') \right| \\
&\leq \frac{1}{\rho^{-1}-1} \sum_{t' \in \mathbb{N}} \rho^{t'} \|w_1\|_\infty \\
&\leq \|w_1\|_\infty
\end{aligned}$$

Hence $w_1 \in \ell_\infty(\mathbb{Z}, \mathbb{R}) \Rightarrow w_2 \in \ell_\infty(\mathbb{Z}, \mathbb{R})$, and w_2 is well-defined by (6). Now substitute (6) in (7) and verify that (w_1, w_2) satisfy (7).

To show the converse, assume that $w_1, w_2 \in \ell_\infty(\mathbb{Z}, \mathbb{R})$ satisfy (7). We need to show that it satisfies (6). Define w_2' by (7) (with w_1 fixed). Then as we have just proven $w_2' \in \ell_\infty(\mathbb{Z}, \mathbb{R})$ and satisfies (7). Hence $w_2 - w_2' \in \ell_\infty(\mathbb{Z}, \mathbb{R})$ and, since (7) defines a linear system, $(0, w_2 - w_2')$ satisfies (7). Let $\Delta := w_2 - w_2'$. Then Δ satisfies $\Delta(t) = \rho\Delta(t-1)$, i.e., $\Delta(t) = \rho^t \Delta(0)$. Since Δ is bounded (on $\mathbb{R}$!), this implies $\Delta = 0$. Hence, $w_2 = w_2'$, and (w_1, w_2) satisfies (6).

What we have used here is that while (7) has many solutions for each w_1, it has only one bounded solution if w_1 is bounded. It is this solution that is given by (6).

(b) Compare the computational complexity of (3), (6), and (7).

Per time step, (3)requires $T-1$ additions and one multiplication, (6)requires in principle an infinite number of multiplications, and (7) requires one addition and two multiplications. Exponential weighting implemented by (7) is hence for several reasons to be preferred above (MA) systems.

(c) Is (7) controllable?

Then R corresponding to (7) is

$$\begin{bmatrix} -\frac{\rho}{\rho^{-1}-1}\xi^{-1} & 1 - \frac{1}{\rho^{-1}-1}\xi^{-1} \end{bmatrix}.$$

There is no common factor, so that the system is controllable.

Obviously, these are reasons enough to prefer exponential weighting implemented by (7) over Moving Average for data smoothing.