Feedback control systems

methods and technologies for mechatronic engineers

Andrea Maria Zanchettin
Professor at Politecnico di Milano

Self Publishers Worldwide

This book was typeset using LaTeX software.

ISBN: 9791221085259

Table of Contents

Chapter 1

Introduction

An automotive company produces and commercialises trucks. In order to stay competitive within its market, the research and development engineering team has been requested to implement a cruise control functionality to maintain constant the speed of the vehicle during long trips.

Figure 1.1: A motivating example: cruise control for a truck.

One of the research engineers had the idea to perform experiments to characterise the behaviour of the vehicle in terms on how the steady-state velocity y depends on the position of the throttle valve u. Then, the data scientist of the research team provided an optimal model to fit the data in the form of a function $f(\cdot)$ such that

$$y = f(u)$$

minimises the prediction error evaluated on the dataset (see Fig. 1.2). With no further ideas, the team decided to use the tool provided by the data scientist to implement the following simple law

$$u = f^{-1}\left(y^0\right)$$

which computes and then actuates the aperture of the throttle valve u based on the desired speed y^0 which, in turn, is set by the driver. Sure regarding the effectiveness of this simple strategy and confident the approach will surely work since

$$y = f(u) = f\left(f^{-1}\left(y^0\right)\right) = y^0$$

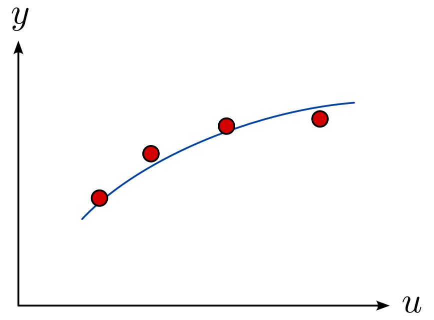


Figure 1.2: Collected experimental data (red dots) and corresponding curve fitting (blue).

the team immediately started the implementation phase and the subsequent deployment for the first tests.
Immediately after the first verification on the field, the team suddenly realised their confidence was premature. The test engineers, in fact, reported the following issues to be solved

- the presence of load on the truck significantly changes the time needed by the truck to reach the steady-state velocity;
- the cruise control system significantly lacks precision on slopes, in particular it reaches values lower than y^0 in uphill conditions, and higher downhill.

With a reasonable disappointment, the research and development team started a month of overnight work to solve the issues. They performed many experiments, in different driving conditions (wheather, slope of the road, load of the truck), to finally propose a solution still, following the same approach of the previous one, requiring the driver to select the driving modality before activating the cruise control.

The management team was reported on the availability of a working solution but also on its complexity. For this reason, the managers decided to setup a **second team** to develop an alternative solution. In few weeks, the second team presented a very simple algorithm that worked as follows

- if the actual speed of the vehicle y is lower than the desired one y^0, i.e. $y < y^0$, the aperture of the throttle valve u is increased;
- in turn, if the actual speed is higher than the one specified by the driver, i.e. $y > y^0$, the aperture of the throttle valve u is decreased.

What is the main difference between the two approaches?
This book is for those who are eager to find the answer and join the second team.

Part I

Systems theory

Chapter 2

Dynamical systems in state-space

2.1 Why dynamics?

Dynamical systems are mathematical models describing the behaviour of a phenomenon or a process characterised by quantities changing over time. These time-varying quantities are the tangible effects of some causes, thought the complete knowledge of these causes is typically not sufficient to univocally determine how the observed quantities are changing over time. As an example, consider a mass sliding upon a friction-less inclined plane as the one reported in Fig. 2.1.

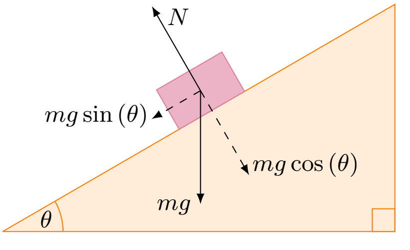


Figure 2.1: A mass sliding on a friction-less inclined surface.

Applying the Newton's second law along the inclined surface, we obtain

$$m\frac{d}{dt}v(t) - mg\sin(\theta) = 0$$

which simplifies in

$$\frac{d}{dt}v(t) = g\sin(\theta) \tag{2.1}$$

Then, the time-varying quantity is clearly the velocity $v()$ of the mass along the inclined surface, while the cause can be clearly identified in the (projection of the) gravitational force mg along the plane. Notice, that the full knowledge about the cause is not sufficient to calculate the velocity of the mass along the plane in an arbitrarily given time instant t. In fact, (2.1) has infinitely many solution which can be written as follows

$$v(t) = \int_0^t g \sin(\theta)\, d\tau = g \sin(\theta)\, t + k$$

Mechanical systems are definitely not the sole characterised by the impossibility to characterise time-varying quantity just from the complete knowledge of the causes making them changing over time. Consider, as an alternative example, the volume $V(t)$ of water within a tank fed with constant flow q. Also in this case

$$V(t) = \int_0^t q d\tau = qt + k$$

In the two given examples, the actual value at time t of the two corresponding time-varying quantities, the velocity $v(t)$ and the volume $V(t)$, was impossible to be determined just knowing the cause letting them to change over the time. On the other hand, the knowledge of them at a certain time instant t_0, i.e. $v(t_0) = v_0$ and $V(t_0) = V_0$, allows one to easily determine both $v(t)$ and $V(t)$ at any time instant t. In fact

$$v(t_0) = v_0 = g \sin(\theta)\, t_0 + k \Rightarrow k = v_0 - g \sin(\theta)\, t_0$$

hence, from the sliding mass

$$v(t) = g \sin(\theta)(t - t_0) + v_0 \tag{2.2}$$

while, equivalently, for the tank

$$V(t_0) = V_0 = qt_0 + k \Rightarrow k = V_0 - qt_0$$

and

$$V(t) = q(t - t_0) + V_0 \tag{2.3}$$

Beside the impressive similarity between the two equations (2.2) and (2.3), notice that in both cases the knowledge of the time-varying quantity in just one specific time instant $t = t_0$, allows one to compute the values of the same quantities in all the others, i.e. $\forall t \neq t_0$. Concluding, we can state that in any dynamical system the values of the time-varying quantities (that we are going to refer to as **states** or **state variables**) at any time instant t can be possibly calculated knowing the causes (gravity and the intake flow, respectively) as well as their values corresponding to a particular time instant t_0.

2.2 Dynamical system

Let $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^n, \boldsymbol{u} \in \mathcal{U} \subseteq \mathbb{R}^{n_u}$ be the (column) vector representing the state of a dynamical system and the (column) vector representing the $n_u \in \mathbb{Z}_{\geq 0}$ exogenous causes (or inputs), respectively. Integer $n \in \mathbb{Z}_{>0}$ (i.e. the number of state variables) is called **order** of the system. Then, the following system of differential equations represents the mathematical model of a (generic) dynamical system:

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} f_1\left(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), u_{n_u}(t), t\right) \\ f_2\left(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), u_{n_u}(t), t\right) \\ \vdots \\ f_n\left(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), u_{n_u}(t), t\right) \end{bmatrix} \tag{2.4}$$

For the sake of compactness, we will make reference to the following (vectorial) form

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t), \boldsymbol{u}(t), t\right) \tag{2.5}$$

with obvious meaning of symbols.

Example 1. *[Mass-spring-damper] Consider a mass m hanged through a spring having stiffness k and a damper with coefficient r which is forced along the vertical direction by an exogenous force f (e.g. gravity) as the one reported in Fig. 2.2.*

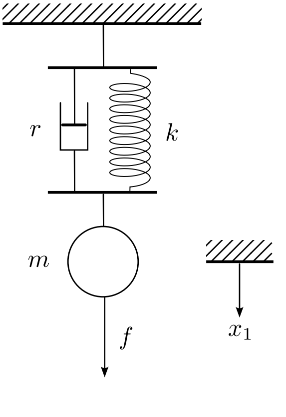


Figure 2.2: A mass-spring-damper system.

Let $x_1 \in \mathbb{R}$ its vertical position and x_2 its vertical velocity, the equations characterising the system are the following ones:

$$\begin{aligned} \frac{d}{dt}x_1(t) &= x_2(t) \\ \frac{d}{dt}x_2(t) &= -\frac{k}{m}x_1(t)\frac{r}{m}x_2(t) + \frac{1}{m}f(t) \end{aligned} \tag{2.6}$$

A mass-spring-damper assembly can be then represented by a second order system ($n = 2$), with $n_u = 1$ input. Notice that the equations characterising this system are linear and there is not explicit dependency on time.

Example 2. *[RL circuit] Consider an electrical circuit consisting in the series of a supplier of voltage v, a resistor with resistance R and an inductor with inductance L. Define x as the current flowing in the circuit, as shown in Figure 2.3.*

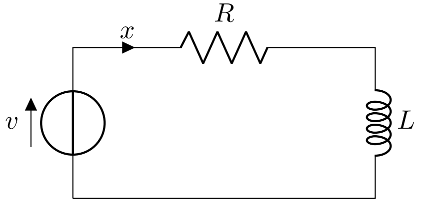


Figure 2.3: A RL electrical circuit.

Applying the Kirchhoff's voltage law, we obtain

$$L\frac{d}{dt}x(t) + Rx(t) = v(t) \tag{2.7}$$

can be simply rewritten as

$$\frac{d}{dt}x(t) = \frac{1}{L}v(t) - \frac{R}{L}x(t) \tag{2.8}$$

The model of the RL circuit contains only $n = 1$ equation and is characterised by $n_u = 1$ inputs. Furthermore, there is no explicit dependency on time t.

Example 3. *[Simple pendulum] Consider a point mass m anchored at the end of a pivoted rod of length L. Let $x_1 \in \mathbb{R}$ its angular displacement with respect to the vertical line and $x_2 \in \mathbb{R}$ its angular velocity as reported in the Figure 2.4.*

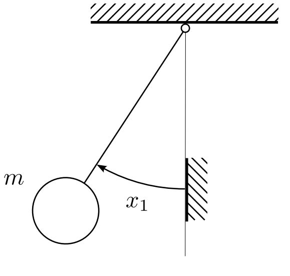


Figure 2.4: A simple pendulum system.

Then, the equations characterising the dynamics of the system are written as follows

$$\begin{aligned}\frac{d}{dt}x_1(t) &= x_2(t)\\ \frac{d}{dt}x_2(t) &= -\frac{g}{L}\sin(x_1(t))\end{aligned} \tag{2.9}$$

The $n = 2$ equations characterising the simple pendulum are non-linear, there is no explicit dependency on time and there are no inputs $n_u = 0$.

2.2.1 State trajectories and their stability

As we have seen, a dynamical system is characterised by differential equations describing how its state variables are evolving with respect to time. Given some knowledge regarding their values at a particular time instant, e.g. $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, we know want to better characterised the evolution of the state variables, i.e. $\boldsymbol{x}(t), \forall t > t_0$.

Definition 1 (Trajectory of the state)**.** *Consider the generic dynamical system in* (2.5) *and assume that state $\boldsymbol{x}(t)$ at time $t = t_0$, i.e. $\boldsymbol{x}(t_0) \in \mathcal{X}$, is known. If the following initial value problem (IVP)*

$$\begin{aligned} &\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t) \\ &\boldsymbol{x}(t_0) \in \mathcal{X} \end{aligned} \tag{2.10}$$

admits a unique solution for $t \geq t_0$, then this solution $\boldsymbol{x}(t)$ is called (state) ***trajectory*** *associated to initial state $\boldsymbol{x}(t_0)$ and input $\boldsymbol{u}(t), t \geq t_0$ (if any).*

Given the trajectory $\boldsymbol{x}(t), t \geq t_0$ associated to initial state $\boldsymbol{x}(t_0)$ and input $\boldsymbol{u}(t), t \geq t_0$, one important aspect to analyse the behaviour of a dynamical system, is to understand how much a **perturbation** on the initial condition $\boldsymbol{x}(t_0)$ will change the corresponding trajectory with respect to $\boldsymbol{x}(t)$. In other words, provided that for initial state $\boldsymbol{x}(t_0)$ and input $\boldsymbol{u}(t), t \geq t_0$ the system (2.5) produces trajectory $\boldsymbol{x}(t)$, what kind of trajectory do we expect for a perturbed initial state $\boldsymbol{x}(t_0) = \tilde{\boldsymbol{x}}(t_0) \in \mathcal{X}$ and for the (same) input $\boldsymbol{u}(t), t \geq t_0$?

Definition 2 (Stability of a trajectory)**.** *A trajectory $\boldsymbol{x}(t), t \geq t_0$ generated from initial state $\boldsymbol{x}(t_0)$ and input $\boldsymbol{u}(t), t \geq t_0$ is called* ***stable*** *if $\forall \varepsilon > 0$ that is "small enough", there exists $\delta > 0$ such that if $\|\boldsymbol{x}(t_0) - \tilde{\boldsymbol{x}}(t_0)\| < \delta$, then $\forall t \geq t_0 : \|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)\| < \varepsilon$.*

According to the definition, a trajectory $\boldsymbol{x}(t)$ corresponding to initial state $\boldsymbol{x}(t_0)$ and input $\boldsymbol{u}(t), t \geq t_0$ is stable when, for the same input, a slightly modified initial condition, i.e. $\|\boldsymbol{x}(t_0) - \tilde{\boldsymbol{x}}(t_0)\| < \delta$, causes a bounded perturbation of the newly generated trajectory, i.e. $\forall t \geq t_0 : \|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)\| < \varepsilon$ (see Fig. 2.5).

As opposed to stability, we introduce the following complementary definition.

Definition 3 (Instability of a trajectory)**.** *A trajectory $\boldsymbol{x}(t), t \geq t_0$ generated from initial state $\boldsymbol{x}(t_0)$ and input $\boldsymbol{u}(t), t \geq t_0$ is called* ***unstable*** *if it is not stable.*

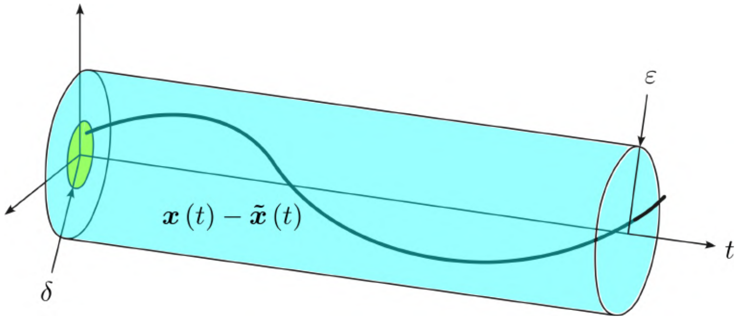


Figure 2.5: Graphical representation of the stability condition.

Finally, in the case of a stable trajectory, it is convenient to further identify the asymptotic ($t \to +\infty$) behaviour of $\|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)\|$.

Definition 4 (Asymptotic stability of a trajectory). *A trajectory $\boldsymbol{x}(t), t \geq t_0$ that has been already proved stable, is called **asymptotically stable** if*

$$\lim_{t \to +\infty} \|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)\| = 0$$

Example 4. *[RL circuit] Consider the state trajectory of the RL circuit when a harmonic input voltage $v(t) = \sin(t), t \geq t_0 = 0$ is applied from a certain initial state $x(0)$. As the equation governing the behaviour of the system is linear and relatively simple, one can easily recognise the corresponding state trajectory being*

$$x(t) = e^{-(R/L)t} x(0) + \frac{Le^{-(R/L)t} - L\cos(t) + R\sin(t)}{L^2 + R^2}, t \geq 0 \tag{2.11}$$

A perturbation on the initial condition will then produce a perturbation on the trajectory which can be evaluated as

$$x(t) - \tilde{x}(t) = e^{-(R/L)t}(x(0) - \tilde{x}(0))$$

showing that the state trajectory $x(t)$ is asymptotically stable.

The notion of asymptotic stability, which further classifies a stable trajectory $\boldsymbol{x}(t)$, simply indicates that the effect of the perturbation on the initial condition $\boldsymbol{x}(t_0) - \tilde{\boldsymbol{x}}(t_0)$ will eventually vanish and, after sufficiently long times, the perturbed trajectory $\tilde{\boldsymbol{x}}(t)$ will look identical to $\boldsymbol{x}(t)$.

It is now convenient to emphasise that the boundedness requirement appearing in the definition of stability is related to the **difference** between trajectory $\boldsymbol{x}(t)$ and the perturbed trajectory $\tilde{\boldsymbol{x}}(t)$, and not to the boundedness of either $\boldsymbol{x}(t)$ or $\tilde{\boldsymbol{x}}(t)$, themselves.

Example 5. *[Unbounded yet stable trajectory] Consider a first-order ($n = 1$) dynamical system that produces, for a certain input $\boldsymbol{u}(t), t \geq t_0$ and a certain initial condition $x_1(0)$, the following state trajectory*

$$x(t) = e^{t-t_0} \sin(t - t_0) + e^{t_0-t} x(t_0), t \geq 0$$

Despite being diverging, trajectory $x(t)$ is (asymptotically) stable. For the same input, a perturbed initial condition $\tilde{x}(t_0)$ will produce the following (perturbed) trajectory:

$$\tilde{x}(t) = e^{t-t_0} \sin(t - t_0) + e^{t_0-t} \tilde{x}(t_0), t \geq t_0$$

*The **difference** between the two can be easily computed as*

$$\begin{aligned} x(t) - \tilde{x}(t) &= e^{t-t_0} \sin(t - t_0) + e^{t_0-t} x(t_0) - \left(e^{t-t_0} \sin(t - t_0) + e^{t_0-t} \tilde{x}_1(t_0)\right) \\ &= e^{t_0-t}(x_1(t_0) - \tilde{x}(t_0)), t \geq t_0 \end{aligned}$$

Hence, in order to make $\|x(t) - \tilde{x}(t)\| < \varepsilon$, one has to just take $0 < \delta < \epsilon$. In fact, if $\|x(t_0) - \tilde{x}(t_0)\| < \delta$, then

$$\|x(t) - \tilde{x}(t)\| = \left\|e^{t_0-t}(x(t_0) - \tilde{x}(t_0))\right\| = e^{t_0-t} \|x(t_0) - \tilde{x}(t_0)\| < \delta < \varepsilon$$

Moreover, it is straightforward to verify its asymptotic stability, in fact

$$\lim_{t\to+\infty} \|x(t) - \tilde{x}(t)\| = \lim_{t\to+\infty} e^{t_0-t} \|x(t_0) - \tilde{x}(t_0)\| = 0$$

2.2.2 Equilibria

Dynamical systems typically have particular state trajectories $\boldsymbol{x}(t), t \geq t_0$ that are **constant** with respect to time, i.e. $\boldsymbol{x}(t) = \bar{\boldsymbol{x}}, \forall t \geq t_0$. More formally

Definition 5 (Equilibrium (constant trajectory))**.** *State trajectory $\boldsymbol{x}(t) = \bar{\boldsymbol{x}}, \forall t \geq t_0$ is called **equilibrium** if $\exists \bar{\boldsymbol{u}} : \boldsymbol{f}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}, t) = \boldsymbol{0}, \forall t \geq t_0$.*

In other words, for constant inputs, i.e. $\boldsymbol{u}(t) = \bar{\boldsymbol{u}}, \forall t \geq t_0$, equilibria are particular (initial) states $\bar{\boldsymbol{x}}$ such that

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= 0, \forall t \geq t_0 \\ \boldsymbol{x}(t_0) &= \bar{\boldsymbol{x}} \in \mathcal{X} \end{aligned} \tag{2.12}$$

Example 6. *[Equilibria of the simple pendulum] The simple pendulum in (2.9) has the following equilibria*

$$x_1(t) = \bar{x}_1 = k\pi, x_2(t) = \bar{x}_2 = 0, k \in \mathbb{Z}$$

which are typically referred to as upward equilibrium (k odd) and downward equilibrium (k even).

Equilibria of a given dynamical system are just special trajectories. Hence, they might be characterised in terms of their stability as shown in Fig. 2.6.

(a) Unstable equilibrium (b) Stable equilibrium

Figure 2.6: Unstable (left) and stable (right) equilibria.

2.2.3 Systems with outputs

As anticipated in the previous Chapter, dynamical systems are typically characterised by of one or more measured variables. To explicitly represent these variables, which are not necessarily state variables, we introduce the output $\boldsymbol{y} \in \mathcal{Y} \subseteq \mathbb{R}^{n_y}$ of a dynamical system as the (column) vector which is related to the state vector, the input vector, and time as follows

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n_y}(t) \end{bmatrix} = \begin{bmatrix} h_1(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), u_{n_u}(t), t) \\ h_2(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), u_{n_u}(t), t) \\ \vdots \\ h_{n_y}(x_1(t), x_2(t), \ldots, x_n(t), u_1(t), u_2(t), u_{n_u}(t), t) \end{bmatrix} \tag{2.13}$$

or, in short:

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{u}(t), t) \tag{2.14}$$

Example 7. *[RL circuit with output] In the RL electrical circuit of Fig. 2.3, one may want to measure the voltage drop on the resistance R. In this case:*

$$y(t) = Rx(t)$$

In this case, the system has $n_y = 1$ outputs.

Example 8. *[Simple pendulum with output] If the mass of the pendulum of Fig. 2.4 is equipped with a laser pointer that is guaranteed to point to the ceiling, one may measure the displacement of its projection with respect to the vertical line. In this case:*

$$y(t) = L\sin(x_1(t))$$

Also in this case, the system has only one output $n_y = 1$.

2.3 LTI dynamical systems

Linear and time-invariant (LTI) systems constitute a particular class of dynamical systems characterised by linear functions $\boldsymbol{f}$ and $\boldsymbol{h}$ with no explicit dependency on time t.

Definition 6 (LTI system)**.** *A LTI system is a particular dynamical system which can be written in the following form*

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t) \end{aligned} \tag{2.15}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times n_u}, \boldsymbol{C} \in \mathbb{R}^{n_y\times n}, \boldsymbol{D} \in \mathbb{R}^{n_y\times n_u}$ are matrices with real coefficients.

Example 9. *[Mass-spring-damper system in matrix form] Consider (again) the mass-spring-damper system and assume its position x_1 to be the output of interest, then its model*

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} f(t) \\ y(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}(t) \end{aligned} \tag{2.16}$$

is a linear and time-invariant (LTI) dynamical system with

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}$$
$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, D = 0$$

2.3.1 Change of variables

Given a generic LTI system as the one in (2.15), it is sometime convenient to change the set of state variables to better highlight some of its properties (we will be more pragmatic on this, later on).

Proposition 1 (Change of variables)**.** *Let $\boldsymbol{T}$ a non-singular constant square matrix of dimension n, then letting $\boldsymbol{z}(t) = \boldsymbol{T}\boldsymbol{x}(t)$, the LTI in (2.15) can be equivalently represented in the following form*

$$\begin{aligned} \frac{d}{dt}\boldsymbol{z}(t) &= \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}\boldsymbol{z}(t) + \boldsymbol{T}\boldsymbol{B}\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{T}^{-1}\boldsymbol{z}(t) + \boldsymbol{D}\boldsymbol{u}(t) \end{aligned} \tag{2.17}$$

Proof. Straightforward by substitution of $\boldsymbol{x}(t)$ with $\boldsymbol{T}^{-1}\boldsymbol{z}(t)$ in (2.15). □

Example 10. *[Change of variables for the mass-spring-damper system] Consider (once again) the mass-spring-damper system and assume the force acting on the ceiling $y = kx_1$ to be the output of interest, then its model is*

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} f(t) \\ y(t) &= \begin{bmatrix} k & 0 \end{bmatrix} \boldsymbol{x}(t) \end{aligned} \tag{2.18}$$

Further assume that the pair x_1, x_2 (position and velocity of the mass) has to be substituted with kx_1, x_2 (force acting on the ceiling and velocity of the mass). Therefore, one can perform the following change of variables

$$\boldsymbol{z}(t) = \begin{bmatrix} kx_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}(t)$$

Hence, according to the new set of state variables, we have $z_1(t) = kx_1(t)$ and $z_2(t) = x_2(t)$. The corresponding equations can be written from (2.6) by computing

$$\begin{aligned}
\frac{d}{dt}z_1(t) &= k\frac{d}{dt}x_1(t) = kx_2(t) = kz_2(t) \\
\frac{d}{dt}z_2(t) &= \frac{d}{dt}x_2(t) = -\frac{k}{m}x_1(t) - \frac{r}{m}x_2(t) + \frac{1}{m}f(t) = \\
&= -\frac{1}{m}z_1(t) - \frac{r}{m}z_2(t) + \frac{1}{m}f(t) \\
y(t) &= kx_1(t) = z_1(t)
\end{aligned} \tag{2.19}$$

It is then left to the reader to prove that

$$\boldsymbol{T}\begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{bmatrix}\boldsymbol{T}^{-1} = \begin{bmatrix} 0 & k \\ -\frac{1}{m} & -\frac{r}{m} \end{bmatrix}$$

and the other identities.

2.3.2 State trajectories for LTI systems and some insights on their stability

The case of LTI systems in the form (2.15) constitutes one of the rare possibilities in which the IVP in (2.10) can be solved in closed-form. In other words, there is a formula to explicitly compute the state trajectories of a given LTI system.

Theorem 1 (Lagrange's formula)**.** *The following IVP*

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \boldsymbol{x}(0) \in \mathcal{X}$$

is solved by the following trajectory

$$\boldsymbol{x}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}(0) + \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\,d\tau, t \geq 0 \tag{2.20}$$

Proof. Left to the reader. □

Example 11. *[State trajectories of the RL circuit] The state trajectory (2.11) of the RL circuit in (2.8) fed with input voltage $v(t) = \sin(t), t \geq 0$ could have been obtained using the Lagrange's formula in (2.20). With $A = -R/L, B = 1/L, u(\tau) = \sin(\tau)$ one obtains:*

$$x(t) = e^{-(R/L)t}x(0) + \int_0^t e^{-(R/L)(t-\tau)}\frac{1}{L}\sin(\tau)\,d\tau =$$

$$= e^{-(R/L)t}x(0) + \frac{1}{L}e^{-(R/L)t}\int_0^t e^{(R/L)\tau}\sin(\tau)\,d\tau =$$
$$= e^{-(R/L)t}x(0) + \frac{1}{L}e^{-(R/L)t}\left[-\frac{e^{(R/L)\tau}\left(L^2\cos(\tau) - RL\sin(\tau)\right)}{R^2+L^2}\right]_0^t =$$
$$= e^{-(R/L)t}x(0) + \frac{1}{L}e^{-(R/L)t}\left(-\frac{e^{(R/L)t}\left(L^2\cos(t) - RL\sin(t)\right)}{R^2+L^2} + \frac{L^2}{R^2+L^2}\right) =$$
$$= e^{-(R/L)t}x(0) + \frac{Le^{-(R/L)t} - L\cos(t) + R\sin(t)}{R^2+L^2}, t \geq 0$$

Along with the obvious benefits of computing closed-form state trajectories for a LTI system, the Lagrange's formula in (2.20) highlights some **fundamental properties** of state trajectories generated by LTI systems.

Property 1. *The state trajectory of a LTI always comprises two terms, i.e.* $e^{\boldsymbol{A}t}\boldsymbol{x}(0)$ *and* $\int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\,d\tau$*, where*

1. *the former (which is call* ***free response****) depends on the initial state* $\boldsymbol{x}(0)$ *only;*

2. *the latter (****forced response****) depends on the input* $\boldsymbol{u}(t)$ *only.*

The free response represents the state trajectory when $\boldsymbol{u}(t) = \mathbf{0}, \forall t \geq 0$*, while the forced response represent the state trajectory when the initial state is the origin of the state space, i.e.* $\boldsymbol{x}(0) = \mathbf{0}$*.*

Property 2 (Superposition principle)**.** *For the linearity of the integral operator, whenever* $\boldsymbol{u}(t) = \boldsymbol{u}_A(t) + \boldsymbol{u}_B(t)$*, the corresponding forced response can be evaluated separately as*

$$\int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\,d\tau = \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}_A(\tau)\,d\tau + \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}_B(\tau)\,d\tau$$

Recalling that stability essentially relates to what happens in case of a perturbation on the initial condition $\boldsymbol{x}(0)$ we can further observe that

Property 3 (Stability of trajectories of LTI systems)**.** *Given the state trajectory* $\boldsymbol{x}(t), t \geq 0$ *of a LTI system, its stability can be determined by the free response, only.*

The given trajectory $\boldsymbol{x}(t), t \geq 0$ is originated by applying input $\boldsymbol{u}(t)$ from the initial state $\boldsymbol{x}(0)$. If the same input is applied from a perturbed initial condition $\tilde{\boldsymbol{x}}(0)$, the perturbed state trajectory can be now evaluated using the Lagrange's formula:

$$\tilde{\boldsymbol{x}}(t) = e^{\boldsymbol{A}t}\tilde{\boldsymbol{x}}(0) + \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\,d\tau, t \geq 0$$

As stability of $\boldsymbol{x}(t)$ depends on the possibility to obtain a bounded perturbation on the state trajectory, i.e. $\|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)\| < \varepsilon$ provided that the perturbation on the initial condition is bounded, i.e. $\|\boldsymbol{x}(0) - \tilde{\boldsymbol{x}}(0)\| < \delta$, one is actually requiring

$$\|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)\| = \left\|e^{\boldsymbol{A}t}\left(\boldsymbol{x}(0) - \tilde{\boldsymbol{x}}(0)\right)\right\| < \varepsilon$$

We have just seen that, for LTI systems, stability of a state trajectory does not depend on the input $\boldsymbol{u}(t)$ but only on the free response. We can now recall another important

Property 4 (Stability of LTI systems)**.** *Given a LTI system, if one state trajectory is (asymptotically) stable, then all the other trajectories are (asymptotically) stable. If one trajectory is unstable, then all the others are unstable.*

In short, for LTI systems stability is a property that is shared among all possible state trajectories generated by arbitrary inputs $\boldsymbol{u}(t)$ from arbitrary initial conditions $\boldsymbol{x}(0)$. For this reason, thought stability has been introduced as property of a state trajectory, in case of LTI systems we can safely introduce the concept of **system stability**.

2.3.3 Stability of LTI systems

We have seen that stability of state trajectories of a LTI, and stability of the system as a whole, can be addressed by analysing whether $\left\|e^{\boldsymbol{A}t}\left(\boldsymbol{x}(0) - \tilde{\boldsymbol{x}}(0)\right)\right\|$ can be made bounded by requiring a proper bound on the perturbation on the initial condition. It should be clear that stability should be strongly related with term $e^{\boldsymbol{A}t}$ and therefore with matrix $\boldsymbol{A}$ which characterises the system.

Stability and eigenvalues

Theorem 2 (Stability analysis of LTI systems)**.** *Depending on matrix $\boldsymbol{A}$, system* (2.15) *is*

- *asymptotically stable if and only if all the eigenvalues of $\boldsymbol{A}$ have strictly negative real parts;*
- *stable (but not asymptotically) if and only if all the eigenvalues of $\boldsymbol{A}$ have non-positive real parts, and those with null real parts are simple roots;*
- *unstable if and only if either at least one of the eigenvalues of $\boldsymbol{A}$ has strictly positive real part, or all the eigenvalues of $\boldsymbol{A}$ have non-positive real parts, and those with null real parts are not simple roots.*

Example 12. *[Stability analysis of a simple mass system] A mass m has position x_1 and velocity x_2 and is pulled by an external force f (see Fig. 2.7). The equations representing the dynamics of this system are as follows*

$$\frac{d}{dt}\boldsymbol{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} f(t) \tag{2.21}$$

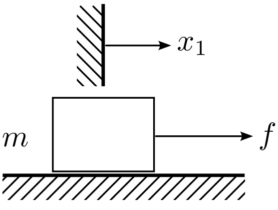


Figure 2.7: A simple mass system.

The eigenvalues of the corresponding matrix $\boldsymbol{A}$ are (complex) solutions of the associated characteristic equation

$$0 = \det\left(\lambda \boldsymbol{I} - \boldsymbol{A}\right)$$

and specifically

$$0 = \det \begin{bmatrix} \lambda & -1 \\ 0 & \lambda \end{bmatrix} = \lambda^2, \bar{\lambda} = 0$$

Hence, matrix $\boldsymbol{A}$ has a zero eigenvalues with algebraic multiplicity 2 and in order to check for stability, eigenvectors have to be computed as well, by solving for $\boldsymbol{v}$ the following equation

$$\boldsymbol{A}\boldsymbol{v} = \bar{\lambda}\boldsymbol{v}$$

Eigenvectors of $\boldsymbol{A}$ associated to $\bar{\lambda} = 0$ are of the form $\boldsymbol{v} = k\begin{bmatrix}1 & 0\end{bmatrix}^T, k \neq 0$, hence the geometric multiplicity is 1. Being different from the algebraic multiplicity, we can conclude that $\bar{\lambda} = 0$ is not a simple root and the system is unstable. If this conclusion seems counter-intuitive, consider the (unpacked) equations of the system

$$\begin{aligned} \frac{d}{dt}x_1(t) &= x_2(t) \\ \frac{d}{dt}x_2(t) &= \frac{1}{m}f(t) \end{aligned} \tag{2.22}$$

From an arbitrary initial state $\boldsymbol{x}(0)$, the free response of the system ($f(t) = 0, \forall t \geq 0$) can be written as follows

$$\begin{aligned} x_2(t) &= x_2(0) \\ x_1(t) &= x_2(0)\,t + x_1(0) \end{aligned} \tag{2.23}$$

from which it should now result evident that the effect of a perturbation on $x_2(0)$ cannot be made bounded.

Finally, consider the LTI system in (2.15) and the same system (2.17) written in terms of a new vector of state variables $\boldsymbol{z}(t) = \boldsymbol{T}\boldsymbol{x}(t), \det \boldsymbol{T} \neq 0$. Notice that the two systems are characterised by matrices $\boldsymbol{A}$ and $\boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}$ which are similar to each other and hence have the same eigenvalues.

Property 5 (Invariance of stability)**.** *The LTI system in* (2.15) *is (asymptotically) stable/unstable if and only if the LTI system is* (2.17) *is also (asymptotically) stable/unstable. That is: a change of variable $\boldsymbol{z}(t) = \boldsymbol{T}\boldsymbol{x}(t), \det \boldsymbol{T} \neq 0$ does not effect the stability property of a LTI system.*

The Routh-Hurwitz criterion

Despite the availability of numerical tools, sometimes it is not possible to evaluate the eigenvalues of matrix $\boldsymbol{A}$ either because of numerical problems, or simply because the matrix is not numerical as it depends on one or more parameters. In this case, an alternative way to evaluate the asymptotic stability of an LTI system is represented by the Routh-Hurwitz criterion.

Definition 7 (Routh-Hurwitz table)**.** *Given the characteristic polynomial associated to a square matrix $\boldsymbol{A}$ of size n, written in the following form*

$$\det\left(\lambda \boldsymbol{I} - \boldsymbol{A}\right) = \alpha_0\lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$$

the Routh-Hurwitz table contains $n+1$ rows and looks like the following one

$$\begin{array}{ccccc}
\alpha_0 & \alpha_2 & \dots & 0 & \dots \\
\alpha_1 & \alpha_3 & \dots & 0 & \dots \\
\vdots & \vdots & & & \\
h_1 & h_2 & \dots & & \\
k_1 & k_2 & \dots & & \\
l_1 & l_2 & \dots & & \\
\vdots & & & &
\end{array}$$

where the first two rows contain the coefficients of the characteristic polynomial (filled one column at a time starting from the coefficient of the highest degree), while the others are computed recursively as

$$l_i = -\frac{1}{k_1}\det\begin{bmatrix} h_1 & h_{i+1} \\ k_1 & k_{i+1} \end{bmatrix}$$

Theorem 3 (Routh-Hurwitz criterion)**.** *All the eigenvalues of matrix $\boldsymbol{A}$ have strictly negative real parts if and only if all the elements within the first column of the Routh-Hurwitz table are either positive or negative.*

The Routh-Hurwitz criterion allows one to address the asymptotic stability property of a given LTI system without explicitly computing the eigenvalues, being just based on the coefficients of the characteristic polynomial.

Example 13. *[Parametric third order system] Consider the following matrix*

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -\beta & -\alpha \end{bmatrix}$$

its characteristic polynomial can be computed as

$$\det\left(\lambda \boldsymbol{I} - \boldsymbol{A}\right) = \det\begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 1 & \beta & \lambda+\alpha \end{bmatrix} = \lambda^3 + \alpha\lambda^2 + \beta\lambda + 1$$

and the corresponding Routh-Hurwitz table is

$$
\begin{array}{ccc}
1 & \beta & 0 \\
\alpha & 1 & 0 \\
\dfrac{\alpha\beta - 1}{\alpha} & 0 & \\
1 & &
\end{array}
$$

For the asymptotic stability of the corresponding LTI we need all elements of the first column to be strictly positive, hence the system is asymptotically stable if and only if

$$
\begin{cases}
\alpha > 0 \\
\alpha\beta > 1
\end{cases}
$$

The corresponding region in the $\alpha\beta$-plane is reported in Fig. 2.8.

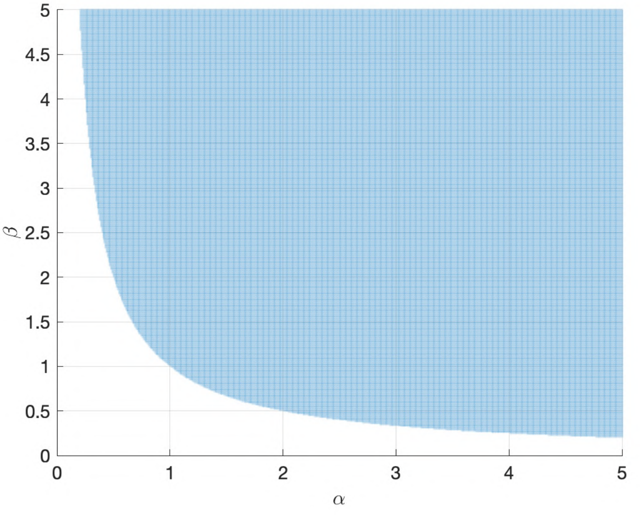


Figure 2.8: Region of the $\alpha\beta$-plane for which the LTI system in this example results asymptotically stable.

2.3.4 Asymptotically stable LTI systems with inputs

The asymptotic stability property also affects the characteristics of the overall response of a LTI system. In particular

Property 6. *Consider the LTI system in* (2.15) *and assume it is asymptotically stable. Then, the following properties hold:*

1. *for a constant input* $\boldsymbol{u}(t) = \bar{\boldsymbol{u}}, \forall t \geq 0$*, the equilibrium is unique and characterised by*
$$\bar{\boldsymbol{x}} = -\boldsymbol{A}^{-1}\boldsymbol{B}\bar{\boldsymbol{u}}$$

2. *after a sufficiently long time, the trajectory of the state will depend only on the input, as the free response eventually vanishes;*

3. *if the input* $\boldsymbol{u}(t)$ *tends to a constant value* $\bar{\boldsymbol{u}}$*, then the trajectory of the state tends to the corresponding equilibrium* $\bar{\boldsymbol{x}} = -\boldsymbol{A}^{-1}\boldsymbol{B}\bar{\boldsymbol{u}}$*.*

The aforementioned properties allow one to study the asymptotic behaviour the the state trajectory, i.e. $\lim_{t\to+\infty}\boldsymbol{x}(t)$ without explicitly applying the Lagrange's formula.

Example 14. *[Asymptotic behaviour of the RL circuit] The state trajectory* (2.11) *of the RL circuit in* (2.8) *fed with a constant input voltage* $v(t) = \bar{v}, \forall t \geq 0$ *could be obviously obtained using the Lagrange's formula in* (2.20)*. With* $A = -R/L, B = 1/L, u(\tau) = \bar{v}, \forall\tau$ *one obtains:*

$$\begin{aligned}
x(t) &= e^{-(R/L)t}x(0) + \bar{v}\frac{1}{L}\int_0^t e^{-(R/L)(t-\tau)}d\tau = \\
&= e^{-(R/L)t}x(0) + \bar{v}\frac{1}{L}e^{-(R/L)t}\int_0^t e^{(R/L)\tau}d\tau = \\
&= e^{-(R/L)t}x(0) + \bar{v}\frac{1}{R}e^{-(R/L)t}\left(e^{(R/L)t} - 1\right) = \\
&= e^{-(R/L)t}x(0) + \bar{v}\frac{1}{R}\left(1 - e^{-(R/L)t}\right)
\end{aligned}$$

It is straightforward to verify that

$$\lim_{t\to+\infty} x(t) = \bar{v}\frac{1}{R} = -\left(\frac{R}{L}\right)^{-1}\frac{1}{L}\bar{v}$$

2.4 Reachability and observability of LTI systems

Having in mind the problem of designing a feedback controller to govern the behaviour of the LTI system in (2.15), there are two fundamental questions to address:

1. how much can we modify the evolution of its state $\boldsymbol{x}(t)$ (and consequently the one of the output, $\boldsymbol{y}(t)$) by designing a proper input $\boldsymbol{u}(t)$?

2. how much information can we infer about the evolution of its state $\boldsymbol{x}(t)$ from the evolution of the output $\boldsymbol{y}(t)$ that we measure?

2.4.1 Reachability

The first question can be addressed by introducing the following

Definition 8 (Reachability)**.** *The LTI system in* (2.15) *is* ***reachable*** *if, for any initial state* $\boldsymbol{x}(0)$*, there exists an input function* $\boldsymbol{u}(\cdot)$ *able to steer the state to the origin in finite time, that is*

$$\forall \boldsymbol{x}(0) : \exists \boldsymbol{u}(\cdot), T : \boldsymbol{x}(T) = \boldsymbol{0}$$

We have now formalized the concept of reachability as the possibility to change the state of the system (actually the possibility to bring the state to the origin for any arbitrary initial condition). This is very important for control purposes as it guarantees we can change the state of a system acting on its inputs (which is indeed the final aim of a control system). On the other hand, the definition itself cannot be used in practice for the verification of such a property.

We then introduce the following

Theorem 4 (Kalman's reachability test)**.** *The LTI system* (2.15) *is reachable if and only if the following matrix (reachability matrix)*

$$\boldsymbol{K_R} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{AB} & \dots & \boldsymbol{A}^{n-1}\boldsymbol{B} \end{bmatrix}$$

is full rank.

which constitutes an easy way to check for the reachability property of (2.15).

Example 15. *[RLL circuit] Consider an electrical circuit consisting in the parallel of a supplier of current* i*, a resistor with resistance* R *and two inductors both with inductance* L*. The voltage drop* v *on the resistor is assumed as output. Define* x_1 *as the current flowing in the first inductor and* x_2 *the one flowing in the second, as shown in Figure 2.9.*

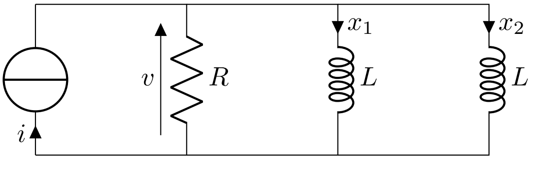


Figure 2.9: A RLL electrical circuit.

The equations characterising the behaviour of this electrical system can be written as follows

$$\begin{aligned}
\frac{d}{dt}x_1(t) &= -\frac{R}{L}x_1(t) - \frac{R}{L}x_2(t) + \frac{R}{L}i(t) \\
\frac{d}{dt}x_2(t) &= -\frac{R}{L}x_1(t) - \frac{R}{L}x_2(t) + \frac{R}{L}i(t) \\
v(t) &= -Rx_1(t) - Rx_2(t) + Ri(t)
\end{aligned} \tag{2.24}$$

which corresponds to the following matrices

$$\boldsymbol{A} = -\frac{R}{L}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}, \boldsymbol{B} = \frac{R}{L}\begin{bmatrix}1\\ 1\end{bmatrix}, \boldsymbol{C} = -R\begin{bmatrix}1 & 1\end{bmatrix}, D = R$$

The corresponding reachability matrix

$$\boldsymbol{K_R} = \begin{bmatrix}\boldsymbol{B} & \boldsymbol{AB}\end{bmatrix} = \frac{R}{L}\begin{bmatrix}1 & -2\frac{R}{L}\\ 1 & -2\frac{R}{L}\end{bmatrix}$$

cannot be full rank as it possesses two linearly dependent (actually identical) rows. Hence the system is non-reachable.

Consider the LTI system in (2.15) and the same system (2.17) written in terms of a new vector of state variables $\boldsymbol{z}(t) = \boldsymbol{Tx}(t), \det \boldsymbol{T} \neq 0$. Notice that the two systems are characterised by matrices $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{TAT}^{-1}, \boldsymbol{TB}$, the corresponding reachability matrices are $\boldsymbol{K_R}$ and

$$\begin{bmatrix}\boldsymbol{TB} & \boldsymbol{TAT}^{-1}\boldsymbol{TB} & \dots & \boldsymbol{TA}^{n-1}\boldsymbol{B}\end{bmatrix} = \boldsymbol{TK_R}$$

respectively. Hence, since $\boldsymbol{T}$ is non-singular, their rank is the same.

Property 7 (Invariance of reachability)**.** *The LTI system in* (2.15) *is reachable if and only if the LTI system is* (2.17) *is also reachable. That is: a change of variable* $\boldsymbol{z}(t) = \boldsymbol{Tx}(t), \det \boldsymbol{T} \neq 0$ *does not effect the reachability property of a LTI system.*

Example 16. *[RLL circuit - continued] Back to the example of the RLL circuit, it might be interesting to better understand the reason of the non-reachability, that is why it is not possible to find a suitable supply current* $i(\cdot)$ *to bring the two currents* x_1 *and* x_2 *to zero in finite time.*
The previous result on the invariance of the reachability property with respect to a change of variable $\boldsymbol{z}(t) = \boldsymbol{Tx}(t), \det \boldsymbol{T} \neq 0$ *is surely helpful. Then, apply the following change of variable*

$$z_1(t) = x_1(t) + x_2(t), z_2(t) = x_1(t) - x_2(t)$$

which corresponds to

$$\boldsymbol{T} = \begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}, \det \boldsymbol{T} \neq 0$$

The same system, written with respect to the newly defined state variable $\boldsymbol{z}(t)$ *corresponds to the following equations*

$$\begin{aligned}&\frac{d}{dt}z_1(t) = -2\frac{R}{L}z_1(t) + 2\frac{R}{L}i(t)\\ &\frac{d}{dt}z_2(t) = 0\end{aligned} \tag{2.25}$$

which show that the supplied current is surely able to modify state z_1*, while* z_2 *evolves regardless the input current* i*. It should be clear that for all initial conditions with* $z_2(0) \neq 0$*, it is not possible to steer the state of the system to the origin in finite time through a suitable input.*
The second state variable is actually not influenced by the input current i *neither directly (in*

the corresponding equation), nor indirectly through z_1.
On the other hand, no matter what $i(t)$ is, state variable z_1 will asymptotically reach the origin, though not in finite time. In fact, applying the Lagrange's formula to the second equation, one obtains

$$x_2(t) = e^{-t} x_2(0), t \geq 0 \tag{2.26}$$

Example 17. *[A three-mass system] Consider a system with three masses connected by springs as the one reported in Fig. 2.10.*

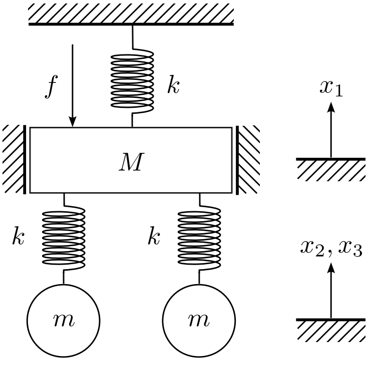


Figure 2.10: A three-mass system.

The differential equations characterising the behaviour of this system can be simply written as follows

$$\begin{aligned}
\frac{d}{dt}v_1(t) &= -\frac{k}{M}x_1(t) - \frac{k}{M}(x_1(t) - x_2(t)) - \frac{k}{M}f(t) \\
\frac{d}{dt}v_2(t) &= -\frac{k}{m}(x_2(t) - x_1(t)) \\
\frac{d}{dt}v_3(t) &= -\frac{k}{m}(x_3(t) - x_1(t))
\end{aligned} \tag{2.27}$$

together with

$$\frac{d}{dt}x_i(t) = v_i(t), i = 1, 2, 3$$

The matrices characterising the LTI systems are given in the following:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{3k}{M} & 0 & \frac{k}{M} & 0 & \frac{k}{M} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{k}{m} & 0 & -\frac{k}{m} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{k}{m} & 0 & 0 & 0 & -\frac{k}{m} & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 0 \\ -\frac{1}{M} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

For any value of the parameters M, m and k the system is non-reachable, in fact

$$\det(\boldsymbol{K_R}) = 0$$

The reason of the lack of reachability can be simply verified by considering

$$p(t) = x_2(t) - x_3(t), v(t) = \frac{d}{dt}p(t) = v_2(t) - v_3(t)$$

Its time-derivative, in fact, results

$$\frac{d}{dt}v(t) = \frac{d}{dt}v_2(t) - \frac{d}{dt}v_3(t) = -\frac{k}{m}x_2(t) + \frac{k}{m}x_3(t) = -\frac{k}{m}p(t)$$

Therefore, a part of the system can be equivalently characterised by equations

$$\begin{aligned} \frac{d}{dt}p(t) &= v(t) \\ \frac{d}{dt}v(t) &= -\frac{k}{m}p(t) \end{aligned} \tag{2.28}$$

which are not influenced by the input force $f(t)$.

2.4.2 Observability

The second question was about the possibility to infer something about the evolution of the state $\boldsymbol{x}(t)$ of (2.15) looking at the outputs $\boldsymbol{y}(t)$. This question can be addressed by introducing the following

Definition 9 (Observability)**.** *The LTI system in* (2.15) *is* ***observable*** *if, for any non null initial condition, the free response of the output differs, within any time horizon, from the zero output, that is*

$$\forall \boldsymbol{x}(0) \neq 0, \forall [t_1, t_2], t_2 > t_1 \geq 0 : \forall t \in [t_1, t_2]\, \boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{x}(0) \neq \boldsymbol{0}$$

Beside control problems, observability is a nice-to-have property, e.g. in fault detection, as a null output will surely imply a null state, i.e. no alarms (outputs are all mute) would surely imply no faulty states (zero state). As done for reachability, we introduce the

Theorem 5 (Kalman's observability test)**.** *The LTI system* (2.15) *is observable if and only if the following matrix (observability matrix)*

$$\boldsymbol{K_O} = \begin{bmatrix} \boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T & \dots & \left(\boldsymbol{A}^T\right)^{n-1}\boldsymbol{C}^T \end{bmatrix}$$

is full rank.

which constitutes an easy way to check for the observability property of (2.15). Moreover, as for the stability and reachability properties, invariance also holds for observability:

Property 8 (Invariance of observability)**.** *The LTI system in* (2.15) *is observable if and only if the LTI system is* (2.17) *is also observable. That is: a change of variable* $\boldsymbol{z}(t) = \boldsymbol{T}\boldsymbol{x}(t), \det \boldsymbol{T} \neq 0$ *does not effect the observability property of a LTI system.*

Example 18. *[RLL circuit - continued] In the example of the RLL circuit, the observability matrix results*

$$\boldsymbol{K_O} = \begin{bmatrix} \boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T \end{bmatrix} = R \begin{bmatrix} -1 & 2\dfrac{R}{L} \\ -1 & 2\dfrac{R}{L} \end{bmatrix}$$

cannot be full rank as it possesses two linearly dependent (actually identical) rows. Hence the system is non-observable. Including the output equation, the same system, written with respect to the state variable $\boldsymbol{z}(t)$ corresponds to the following equations

$$\begin{aligned} \frac{d}{dt}z_1(t) &= -2\frac{R}{L}z_1(t) + 2\frac{R}{L}i(t) \\ \frac{d}{dt}z_2(t) &= 0 \\ v(t) &= -Rz_1(t) + Ri(t) \end{aligned} \tag{2.29}$$

If we consider $i(t) = 0$, it should be noted that

$$v(t) = -Re^{-2(R/L)t}z_1(0)$$

Hence for all initial conditions with $z_1(0) = 0, z_2(0) \neq 0$, the output voltage $v(t)$ would be zero, despite the non-zero initial condition.
It should be noted that the behaviour of the output voltage $v(t)$ is not influenced by the behaviour of the second state variable $z_2(t)$ neither directly (in the output equation), nor indirectly, through the behaviour of $z_1(t)$.

2.4.3 Kalman's decomposition

Given a generic LTI system like the one in (2.15), it is always possible to find a change of variables $\boldsymbol{z}(t) = \boldsymbol{T}\boldsymbol{x}(t)$, $\det \boldsymbol{T} \neq 0$ that allows the equations of the system to be clustered in four (possibly empty) blocks: reachable and non-reachable, observable and non-observable. This situation, which is called **Kalman's decomposition**, is shown in Fig. 2.11 which also reports the possible influences between blocks.

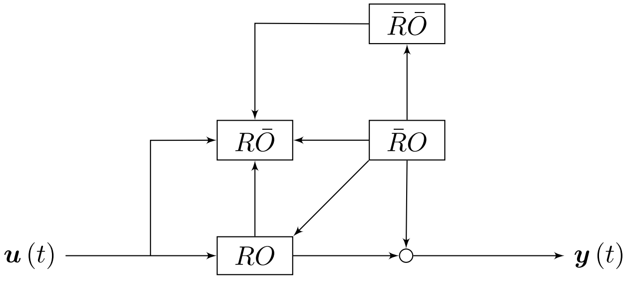


Figure 2.11: Kalman's decomposition of a generic LTI system.

Example 19. *[Kalman's decomposition] Consider the following LTI dynamical system of order $n = 4$, with $n_u = 1$ inputs and $n_y = 1$ outputs:*

$$\begin{aligned}
\frac{d}{dt} z_1(t) &= u(t) \\
\frac{d}{dt} z_2(t) &= z_1(t) + z_3(t) \\
\frac{d}{dt} z_3(t) &= z_3(t) \\
\frac{d}{dt} z_4(t) &= z_3(t) - z_4(t) \\
y(t) &= z_1(t) + z_3(t)
\end{aligned} \tag{2.30}$$

State variable $z_1(t)$ is directly influenced by the input and influences the output. Variable $z_2(t)$ is influenced by the input through $z_1(t)$. Variable $z_3(t)$ is not influenced by the input neither directly nor indirectly, but influences the output. Finally, variable $z_4(t)$ is not influenced by the input nor influences the output variable. Corresponding clusters are reported in Fig. 2.12.

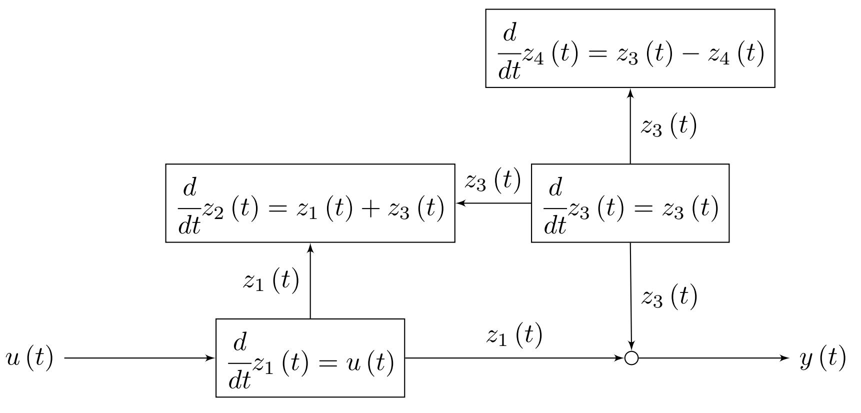


Figure 2.12: Kalman's decomposition of the system in (2.30).

Chapter 3

Transfer function of LTI systems

The analysis of the behaviour of a LTI system is not necessarily limited to its description on a time basis. In all engineering applications, it is often convenient to address the characterisation of a dynamical system in the frequency domain. The reader should be familiar with the concept of (unilateral) Laplace transform of a signal $\boldsymbol{v}(t), t \geq 0$ which is the quantity defined as follows

$$\mathcal{L}[\boldsymbol{v}(t)] = \boldsymbol{V}(s) = \int_0^{+\infty} e^{-st}\boldsymbol{v}(t)\,dt, s \in \mathbb{C}$$

In the following, we will understand how the concept of the Laplace transform can be applied to signals that are typically characterising LTI systems, i.e. its inputs $\boldsymbol{u}(t)$, its state $\boldsymbol{x}(t)$, and its outputs $\boldsymbol{y}(t)$.

3.1 Transfer function

Consider a generic LTI system as the one in (2.15) and apply the Laplace transform to $\boldsymbol{u}(t)$, $\boldsymbol{x}(t)$, and $\boldsymbol{y}(t)$

$$\mathcal{L}[\boldsymbol{u}(t)] = \boldsymbol{U}(s), \mathcal{L}[\boldsymbol{x}(t)] = \boldsymbol{X}(s), \mathcal{L}[\boldsymbol{y}(t)] = \boldsymbol{Y}(s)$$

and

$$\mathcal{L}\left[\frac{d}{dt}\boldsymbol{x}(t)\right] = s\boldsymbol{X}(s) - \boldsymbol{x}(0)$$

By substituting in the equation of the system, one obtains

$$s\boldsymbol{X}(s) - \boldsymbol{x}(0) = \boldsymbol{A}\boldsymbol{X}(s) + \boldsymbol{B}\boldsymbol{U}(s)$$

$$\boldsymbol{Y}(s) = \boldsymbol{C}\boldsymbol{X}(s) + \boldsymbol{D}\boldsymbol{U}(s)$$

Rearranging the two previous equations, one obtains

$$\boldsymbol{Y}(s) = \left[\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}\right]\boldsymbol{U}(s)+\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{x}(0)$$

From the Lagrange's formula we already know that

$$\boldsymbol{y}(t) = \boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{x}(0)+\boldsymbol{C}\int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\,d\tau+\boldsymbol{D}\boldsymbol{u}(t),t\geq 0$$

hence

$$\mathcal{L}\left[\boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{x}(0)\right] = \boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{x}(0)$$

$$\mathcal{L}\left[\boldsymbol{C}\int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)\,d\tau+\boldsymbol{D}\boldsymbol{u}(t)\right] = \left[\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}\right]\boldsymbol{U}(s)$$

the two terms represent the Laplace transforms of the free and the forced responses of the output $\boldsymbol{y}(t)$, respectively.

Definition 10 (Transfer function)**.** *The quantity*

$$\boldsymbol{G}(s) = \left[\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}\right] \tag{3.1}$$

is called ***transfer function*** *of system* (2.15) *and represents the relationship between the Laplace transform of the input* $\boldsymbol{U}(s)$ *and the Laplace transform of the forced response of the output* $\boldsymbol{Y^f}(s)$*, so that*

$$\boldsymbol{Y^f}(s) = \boldsymbol{G}(s)\boldsymbol{U}(s) \tag{3.2}$$

Example 20. *[Transfer function of the RL circuit] The RL circuit in Fig. 2.3 having* $y(t) = Rx(t)$ *as output equation has the following transfer function*

$$G(s) = R\left(s+\frac{R}{L}\right)^{-1}\frac{1}{L} = \frac{R/L}{s+R/L} \tag{3.3}$$

To verify the property of the transfer function being the relationship between the Laplace transform of the input and the Laplace transform of the forced response of the output, consider a step signal as input voltage for the circuit, i.e. $v(t) = 1, t \geq 0$*. Using the Lagrange's formula (for* $x(0) = 0$*) one obtains*

$$y^f(t) = \frac{R}{L}\int_0^t e^{-(R/L)(t-\tau)}d\tau = \frac{R}{L}e^{-(R/L)t}\int_0^t e^{-(R/L)\tau}d\tau = 1-e^{-(R/L)t}, t\geq 0$$

On the other hand

$$Y^f(s) = G(s)V(s) = \frac{R/L}{(s+R/L)\,s} = \frac{1}{s}-\frac{1}{s+R/L}$$

which indeed confirms that $\mathcal{L}\left[y^f(t)\right] = Y^f(s) = G(s)V(s)$*.*

3.2 Transfer function of SISO systems

In the following, we will confine our analyses to a subset of LTI systems which are characterised by exactly one input and one output, i.e.

Definition 11 (SISO system)**.** *A single-input single-output (SISO) LTI system is a LTI with exactly $n_u = 1$ input and exactly $n_y = 1$ output, i.e. a LTI of the kind*

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}u(t) \\ y(t) &= \boldsymbol{C}\boldsymbol{x}(t) + Du(t) \end{aligned} \tag{3.4}$$

The transfer function of a SISO system is therefore the scalar quantity

$$G(s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + D \tag{3.5}$$

and can be rewritten as follows

$$G(s) = \frac{\boldsymbol{C}\text{adj}(s\boldsymbol{I} - \boldsymbol{A})\boldsymbol{B}}{\det(s\boldsymbol{I} - \boldsymbol{A})} + D = \frac{\boldsymbol{C}\text{adj}(s\boldsymbol{I} - \boldsymbol{A})\boldsymbol{B} + \det(s\boldsymbol{I} - \boldsymbol{A})D}{\det(s\boldsymbol{I} - \boldsymbol{A})}$$

Apart from cancellation between terms appearing both in the numerator and in the denominator, the denominator of the transfer function of a SISO system $\det(s\boldsymbol{I} - \boldsymbol{A})$ is the characteristic polynomial of matrix $\boldsymbol{A}$, while the numerator is the sum of the characteristic polynomial scaled by D and polynomial $\boldsymbol{C}\text{adj}(s\boldsymbol{I} - \boldsymbol{A})\boldsymbol{B}$. Since the adjugate matrix is obtained by computing minors of $s\boldsymbol{I} - \boldsymbol{A}$ (determinant of matrices obtained by removing one row and one column), the numerator is a polynomial with degree at most equal to the degree of the denominator. The degree of denominator is the order n of the system. More in general:

Property 9. *The transfer function $G(s)$ of a SISO LTI system is the ratio between two polynomials with real coefficients. The degree of the denominator is at most equal to the order n of the system, the degree of the numerator cannot exceed the one of the denominator.*

A generic transfer function associated to a SISO LTI system can be then written in the following form

$$G(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} \tag{3.6}$$

Definition 12. *Given a SISO LTI system and its transfer function $G(s)$ in* (3.6)*:*

1. *a root of the numerator is called* ***zero****;*
2. *a root of the denominator is called* ***pole****;*
3. *the difference between the degree of the denominator and the degree of the numerator $\nu = n - m \geq 0$ is called* ***relative degree****;*
4. *if the relative degree is positive ($\nu > 0$, i.e. $n > m$), which happens whenever $D = 0$, the transfer function is called* ***strictly proper****, otherwise ($n = m$) proper;*
5. *the number $g \in \mathbb{Z}$ of poles ($g > 0$) or zeroes ($g < 0$) in the origin is called* ***type****;*
6. *quantity $\mu = \lim_{s\to 0} s^g G(s)$ is called* ***generalised gain****.*

3.2.1 Stability, reachability, observability

In the following, we will try to understand the relationship between the transfer function, its poles, the eigenvalues of matrix $\boldsymbol{A}$, and the reachability and observability properties.

Example 21. *[Transfer function of the RLL circuit] Consider the transfer function of the RLL electrical circuit in Fig. 2.9:*

$$G(s) = \frac{4R^2 + LRs}{2R + Ls}$$

The transfer function is a first-order one, with a pole in $s = -2R/L$ and one real zero. Since poles are roots of the denominator of the transfer function, they are necessarily also eigenvalues. In fact:

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = s^2 + (2R/L)\,s$$

which nullifies for $\lambda = -2R/L$ and $\lambda = 0$.
The reason why the denominator of the transfer function $G(s)$ has not the same order of the system $n = 2$ is to be further investigated. As of now we might argue this fact to be somehow related to the lack of reachability and/or observability of the system.

Property 10 (Poles are eigenvalues)**.** *Given a LTI SISO system, poles of its transfer function $G(s)$ are also eigenvalues. It is not necessarily true that eigenvalues are also poles.*

Being the transfer function $G(s)$ an operator that relates the input with the forced response of the output, going back to the Kalman's decomposition in Fig. 2.11, it should be clear that the transfer function must contains only the part of the system which is simultaneously reachable and observable.

Property 11. *Given a LTI SISO system, eigenvalues which are not poles, if any, are necessarily related to the parts of the system which are non-reachable or non-observable (or both). In turn, eigenvalues being also poles are those of the reachable and observable part of the system.*

Based on this property, the reader should pay particular attention in discussing the stability property of system (11.3) by looking at the corresponding transfer function (3.5) as the latter does not necessarily contains all the eigenvalues of matrix $\boldsymbol{A}$ as poles. If, in turn, this happens, the corresponding system is surely both reachable and observable.

3.2.2 Representations of a transfer function

The generic form of a SISO transfer function in (3.6) is not suitable for a quick evaluation of the position of poles, zeros, the evaluation of its generalised gain, etc. As an alternative representation, we can introduced the following one

$$G(s) = \frac{\prod_{i=1}^{m_z}(s+z_i)}{\prod_{j=1}^{m_p}(s+p_j)}$$

where the m_p poles and the m_z zeros are immediate from p_i's and z_i's. However, since poles and zeros are solution of polynomial equations, they might be complex and conjugate, a better way to represent a transfer function is the following one

$$G(s) = \frac{\prod_i (s+z_i)}{\prod_j (s+p_j)} \frac{\prod_i \left(s^2 + 2\xi_{z,i}\omega_{z,i}s + \omega_{z,i}^2\right)}{\prod_j \left(s^2 + 2\xi_{p,j}\omega_{p,j}s + \omega_{p,j}^2\right)} \tag{3.7}$$

where p_i's and z_i's are used to identify real poles and zeros, respectively. The second part has been added to represent complex and conjugate poles and zeros, where ω's and ξ's stand for their magnitudes (or natural frequencies) and dampings, respectively.

Another useful representation is the one reported in the following

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1+s\tau_i)}{\prod_j (1+sT_j)} \frac{\prod_i \left(1 + \frac{2\xi_{z,i}}{\omega_{z,i}}s + \frac{s^2}{\omega_{z,i}^2}\right)}{\prod_j \left(1 + \frac{2\xi_{p,j}}{\omega_{p,j}}s + \frac{s^2}{\omega_{p,j}^2}\right)} \tag{3.8}$$

where g and μ are the generalised gain and type, respectively, τ_i's ($\tau_i = 1/z_i$) and T_j's ($T_j = 1/p_j$) are time constants associated to zeros and poles, respectively, and the remaining part is still representing complex and conjugate poles and zeros.

3.3 LTI systems with delayed inputs or outputs

Some LTI systems are characterised by delayed actuations (inputs) and/or a delayed measurements (outputs). Consider, for example, a tank a certain solution has to be processed (see Fig. 3.1). Define $u(t)$ the flow rate of the solution which is governed by another process, not represented. Modelling the quantity

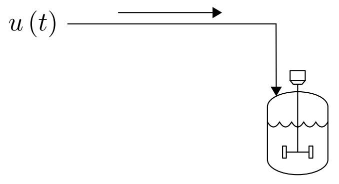


Figure 3.1: Example of a LTI system with a delayed input.

of solution in the tank is straightforward if we define its volume $V(t)$ as state variables. However, the evaluation of its time derivative deserves further discussions. At time t, the flow rate of the solution feeding the tank is not $u(t)$ but rather $u(t-\tau)$ where $\tau \geq 0$ represents the **delay** in traversing the pipe-way and has been assumed to be independent from $u(t)$. As a consequence, the time derivative of the volume $V(t)$ of the solution should be written as

$$\frac{d}{dt}V(t) = u(t-\tau)$$

Overall, the dynamics of the system in Fig. 3.1 can be written as follows

$$\begin{aligned} \frac{d}{dt}V(t) &= u(t-\tau) \\ y(t) &= V(t) \end{aligned} \tag{3.9}$$

For null initial conditions, transforming these equations according to Laplace yields to

$$Y(s) = \frac{e^{-s\tau}}{s}U(s)$$

where

$$\mathcal{L}[u(t-\tau)] = e^{-s\tau}\mathcal{L}[u(t)] = e^{-s\tau}U(s)$$

Note that for the first time, the transfer function $G(s)$ representing the relationship between input and output of a LTI SISO system is no longer a rational function.

Consider the following generic LTI SISO system characterised by a delayed input (similar arguments apply whenever the output is delayed, i.e. $y(t-\tau)$ in place of $y(t)$)

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}u(t-\tau) \\ y(t) &= \boldsymbol{C}\boldsymbol{x}(t) + Du(t-\tau) \end{aligned} \tag{3.10}$$

the corresponding transfer function $G(s)$ can be written as

$$G(s) = \left[\boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + D\right] e^{-s\tau}$$

3.4 Step response of first- and second-order LTI SISO systems

Understanding how the parameters of a SISO transfer function will influence the corresponding forced response is crucial for the design problem that will be introduced in the following Chapters. Moreover, the relation between the parameters of the transfer function and the response of a system is also interesting to quickly derive models of physical systems by just running simple experiments. In the following, we will consider a generic LTI SISO system characterised by transfer function $G(s)$ and we will address how the parameters of such transfer function (e.g. the position of poles and zeros, its generalised gain, etc.) will influence the behaviour of its forced response $y^f(t)$ for a step input $u(t) = 1, t \geq 0$.

Figure 3.2 reports a generic response $y^f(t)$ for a step input applied at $t = 0$. Based on the behaviour of $y^f(t)$, we can define the following quantities

- initial delay $\tau_d \geq 0$;
- steady-state value y^f_∞, i.e. $y^f_\infty = \lim_{t\to+\infty} y^f(t)$, if any;
- rise time T_r, the time interval $T_r = T_{0.9} - T_{0.1}$ such that $y^f(T_{0.1}) = 0.1y^f_\infty$ and $y^f(T_{0.9}) = 0.9y^f_\infty$;
- overshoot S, the difference between the maximum value and the steady-state value, i.e. $S = \sup_t y^t(t) - y^f_\infty$, or equivalently its normalised value

$$S_\% = 100\frac{\sup_t y^f(t) - y^f_\infty}{y^f_\infty}$$

- ε%-settling-time $T_{s\varepsilon}$ the time instant when the response $y^f(t)$ enters and never leaves the bandwidth $y^f_\infty \pm \varepsilon\%$;
- the initial slope of the response, i.e. $\dot{y}^f(\tau_d)$.

In the following, we will analyse how the aforementioned quantities are related to the parameters of transfer function $G(s)$ which will be assumed to be asymptotically stable.

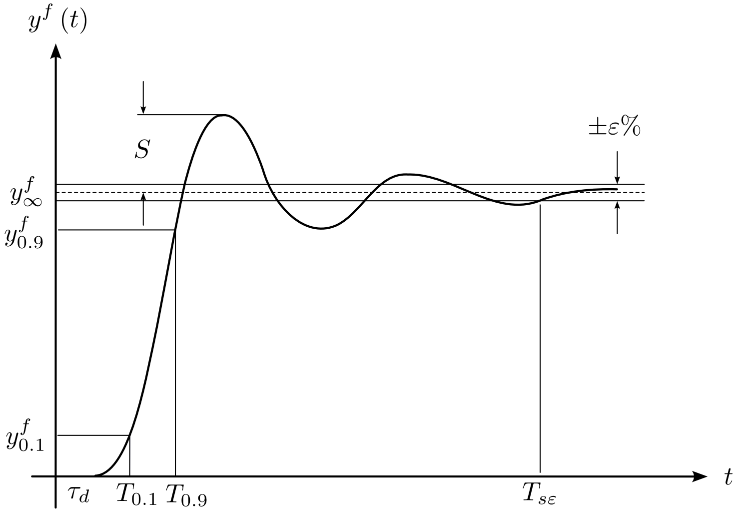


Figure 3.2: Generic forced step response $y^f(t)$ of a LTI SISO system.

3.4.1 Step response of first-order systems

One pole and no zeros

A generic first-order asymptotically stable transfer function $G(s)$ with no zeroes assumes the following form

$$G(s) = \frac{\mu}{1 + sT}, T > 0$$

The forced step response of $G(s)$ results as follows

$$Y^f(s) = G(s)\frac{1}{s} = \frac{\mu}{(1 + sT)s} = \frac{-\mu T}{1 + sT} + \frac{\mu}{s} = \frac{\mu}{s} - \frac{\mu}{s + 1/T}$$

which corresponds, in time domain, to the following

$$y^f(t) = \mu\left(1 - e^{-t/T}\right), t \geq 0 \tag{3.11}$$

By analysing the behaviour of (3.11), we can easily state that

- its steady-state value is $y^f_\infty = \mu$, no matter what T is, in fact

$$y^f_\infty = \lim_{t\to+\infty} y^f(t) = \lim_{s\to 0} sY^f(s) = \mu$$

- the rise time is $T_r = T_{0.9} - T_{0.1} = 2.1972T$, no matter what μ is, in fact

$$x\mu = \mu\left(1 - e^{-T_x/T}\right) \Rightarrow T_x = -T\log_e(1 - x) = \begin{cases} 0.1054T & x = 0.1 \\ 2.3026T & x = 0.9 \end{cases}$$

- the response has no overshoot, i.e. $S = 0$, in fact $\sup_t y^f(t) = y^f_\infty = \mu$;
- the ε%-settling time is $T_{s\varepsilon} = T_{1-\varepsilon} = -T \log_e(\varepsilon)$, and depends on T only.

Additionally, the initial value is zero, i.e. $y^f(0) = 0$, no matter the values of μ and T in fact

$$y^f(0) = \mu\left(1 - e^0\right) = \lim_{s \to +\infty} sY^f(s) = \lim_{s \to +\infty} sY^f(s) = 0$$

in turn, the initial slope is $\dot{y}^f(0) = \mu/T$, in fact

$$\dot{y}^f(0) = \lim_{s \to +\infty} s\left(sY^f(s) - y^f(0)\right) = \lim_{s \to +\infty} s^2 Y^f(s) = \frac{\mu}{T}$$

To summarise, Fig. 3.3 reports the generic forced step response of $G(s)$ in (3.11). It is worth recalling that the response asymptotically approaches the generalised gain μ of the transfer function without overshoot, and the timings only depend on the time constant T, hence on the position of the pole $-1/T$. In particular, the smaller $T > 0$ is (or equivalently, the higher the frequency of the pole $1/|T|$), the faster the response.

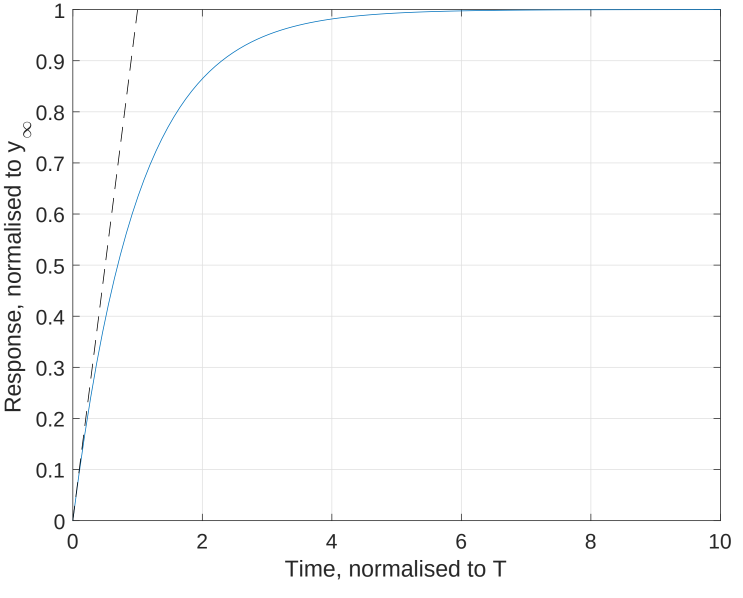


Figure 3.3: Forced step response $y^f(t)$ of a generic asymptotically stable first-order LTI SISO system with no zeros, axes are normalised to the time constant T and to the steady-state value y^f_∞, respectively.

One pole and a zero in the origin

A transfer function of this kind can be written generically as

$$G(s) = \frac{\mu s}{1+sT}, T>0$$

The forced step response of $G(s)$ results as follows

$$Y^f(s) = G(s)\frac{1}{s} = \frac{\mu}{1+sT} = \frac{\mu/T}{s+1/T}$$

which corresponds, in time domain, to the following

$$y^f(t) = -\frac{\mu}{T}e^{-t/T}, t \geq 0 \tag{3.12}$$

Looking at the response in (3.12), we can easily notice that

- the response has a null steady-state value regardless μ and T, in fact
$$y^f_\infty = \lim_{t\to+\infty} y^f(t) = \lim_{s\to 0} sY^f(s) = 0$$
- the ε%-settling time is again $T_{s\varepsilon} = -T\log_e(\varepsilon)$;
- the initial value depends on both μ and T and results $y^f(0) = \mu/T$;
- the initial slope also depends on both μ and T and can be evaluated as

$$\dot{y}^f(0) = \lim_{s\to+\infty} s\left(sY^f(s) - y^f(0)\right) = \lim_{s\to+\infty} s\left(sY^f(s) - \frac{\mu}{T}\right) =$$

$$= \lim_{s\to+\infty} s\left(\frac{s\mu}{1+sT} - \frac{\mu}{T}\right) = \lim_{s\to+\infty} \frac{\mu T s^2 - \mu(1+sT)s}{T(1+sT)} = -\frac{\mu}{T^2} = -\frac{\dot{y}^f(0)}{T}$$

To summarise, Fig. 3.4 reports the generic forced step response of $G(s)$ in (3.12). It is worth to remind that the response monotonically approaches zero, and timings only depends on the time constant T, hence on the position of the pole $-1/T$. In particular, the smaller $T>0$ is (or equivalently, the higher the frequency of the pole $1/|T|$), the faster the response.

3.4.2 Step response of second-order systems

The analysis of second-order system is slightly more involved, as many options have to be addressed. In particular, we will focus on three possible cases for transfer function $G(s)$: two real poles and no zeros, two real poles and a zero (possibly in the origin), and complex and conjugate poles and no zeroes.

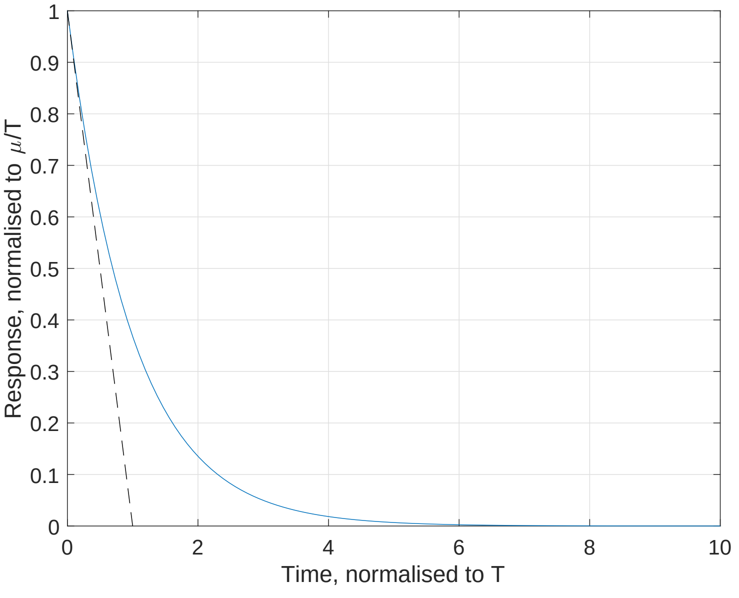


Figure 3.4: Forced step response $y^f(t)$ of a generic asymptotically stable first-order LTI SISO system with a zero in the origin, axes are normalised to the time constant T and to the initial value μ/T, respectively.

Real poles and no zeroes

An asymptotically stable transfer function $G(s)$ of this kind can be generally written as

$$G(s) = \frac{\mu}{(1+sT_1)(1+sT_2)}, T_1 \geq T_2 > 0$$

In the case $T_1 > T_2 > 0$, the forced step response of $G(s)$ results as follows

$$Y^f(s) = G(s)\frac{1}{s} = \frac{\mu}{s(1+sT_1)(1+sT_2)} = \frac{\mu}{s} + \frac{\frac{\mu T_1}{T_1 - T_2}}{1+s/T_1} + \frac{\frac{\mu T_2}{T_1 - T_2}}{1+s/T_2}$$

which corresponds, in time domain, to the following

$$y^f(t) = \mu\left(1 - \frac{T_1}{T_1 - T_2}e^{-t/T_1} - \frac{T_2}{T_1 - T_2}e^{-t/T_2}\right), t \geq 0 \qquad (3.13)$$

from which we can observe that

- its steady-state value is $y^f_\infty = \mu$, no matter what T_1, T_2 are, in fact

$$y^f_\infty = \lim_{t\to+\infty} y^f(t) = \lim_{s\to 0} sY^f(s) = \mu$$

- the response has no overshoot, no matter what μ, T_1, T_2 are, i.e. $S = 0$;
- the initial value is null, no matter what μ and T_1, T_2 are, in fact

$$y^f(0) = \lim_{s \to +\infty} sY^f(s) = \lim_{s \to +\infty} G(s) = 0$$

- the initial slope is also null, no matter what μ and T_1, T_2 are, in fact

$$\dot{y}^f(0) = \lim_{s \to +\infty} s\left(sY^f(s) - y^f(0)\right) = \lim_{s \to +\infty} sG(s) = 0$$

In the case $T_1 \gg T_2 > 0$, the response (3.13) simplifies to

$$y^f(t) = \mu\left(1 - e^{-t/T_1}\right), t \geq 0$$

Hence, whenever the frequency of the two poles $-1/T_1$ and $-1/T_2$ are very well separated, the high-frequency one can be safely neglected and the corresponding forced step response resembles the one of a first-order system with time constant $T = T_1$.

Example 22. *[Settling time with a dominant pole (or time constant)] Consider the following transfer function*

$$G(s) = \frac{7}{(1+3s)(1+60s)}$$

and assume that one wants to evaluate the $\varepsilon\%$-settling time of the corresponding step response with $\varepsilon\% = 1\%$. Then substituting $T_1 = 60, T_2 = 3$ and $\mu = 7$ in (3.13), we obtain

$$y^f(t) = 7\left(1 - \frac{60}{57}e^{-t/60} - \frac{3}{57}e^{-t/3}\right), t \geq 0$$

Since $T_1 = 60 \gg T_2 = 3$, we can also consider the following approximation

$$y^f(t) \approx 7\left(1 - e^{-t/60}\right), t \geq 0$$

The forced step response of the system and its first-order approximation are shown in Fig. 3.5. Given the validity of the approximation, we can conclude that $T_{s1} \approx \log_e(100)T_1 \approx 4.6T_1 = 276$ s.

Finally, the case $T_1 = T_2 = T$ has to be addressed separately. The forced response can be computed as follows

$$Y^f(s) = G(s)\frac{1}{s} = \frac{\mu}{s(1+sT)^2}$$

which corresponds to

$$y^f(t) = \mu\left(1 - e^{-t/T} - \frac{t}{T}e^{-t/T}\right), t \geq 0 \tag{3.14}$$

Observe that:

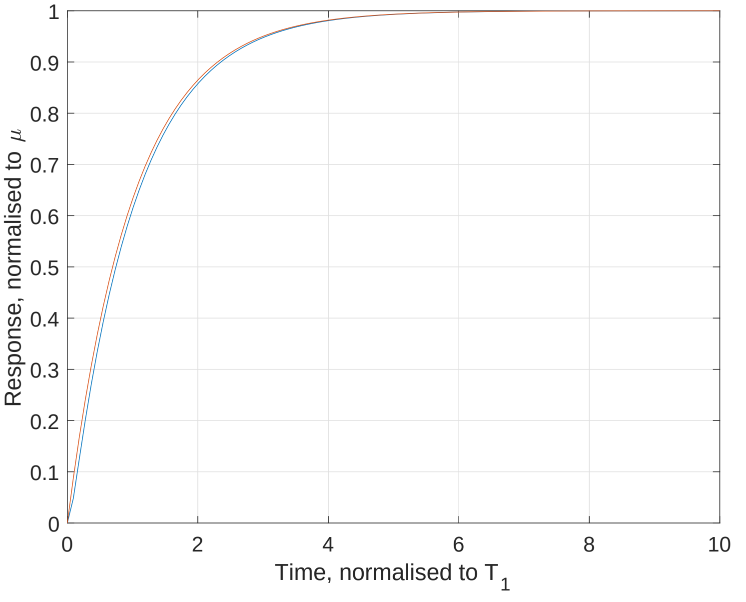


Figure 3.5: Forced step response $y^f(t)$ of a generic asymptotically stable second-order LTI SISO system with a dominant time constant (blue) and its approximating first-order response (red), axes are normalised to the dominant time constant T_1 and to the finale value μ, respectively.

- its steady-state value is $y^f_\infty = \mu$, no matter what T is;
- the response has no overshoot, i.e. $S = 0$, no matter what μ and T are;
- the initial value is null, i.e. $y^f(0) = 0$, regardless the values of μ and T;
- the initial slope is also always null, i.e. $\dot{y}^f(0) = 0$, in fact

$$\dot{y}^f(0) = \lim_{s \to +\infty} s\left(sY^f(s) - y^f(0)\right) = \lim_{s \to +\infty} sG(s) = 0$$

Figure 3.5 reports the generic forced step response for the case $T_2 = T_1 = T$ compared to the one obtained from a first-order system having the same time constant T. As one can see, the second order response is slightly slower than the first-order one having the same time constant.

Real poles and one zero

A second-order asymptotically stable transfer function $G(s)$ with two real poles and one zero can be written as

$$G(s) = \frac{\mu(1 + s\tau)}{(1 + sT_1)(1 + sT_2)}, T_1 \geq T_2 > 0$$

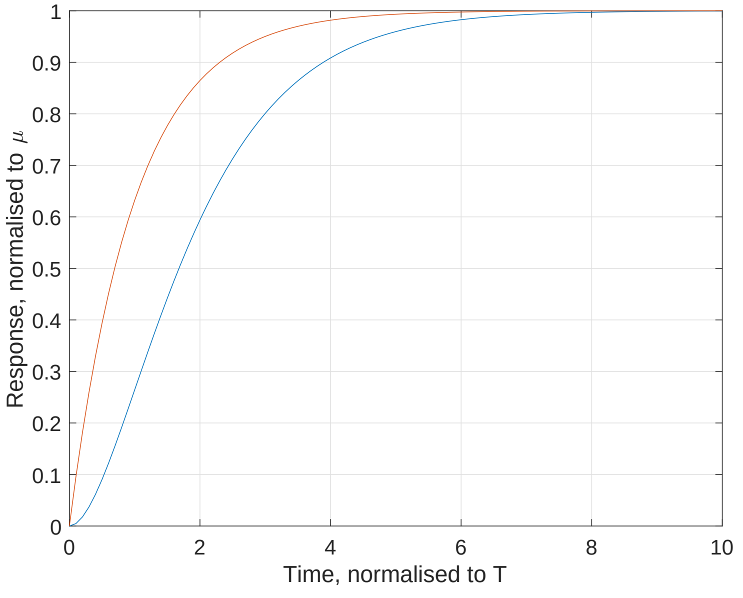


Figure 3.6: Forced step response $y^f(t)$ of a generic asymptotically stable second-order LTI SISO system with $T_1 = T_2 = T$ (blue) as compared with the first-order response (red), axes are normalised to the time constant T and to the steady-state value y^f_∞, respectively.

The Laplace transform of its forced step response results as follows

$$Y^f(s) = G(s)\frac{1}{s} = \frac{\mu(1+s\tau)}{s(1+sT_1)(1+sT_2)}$$

from which we can observe that

- its steady-state value is $y^f_\infty = \mu$, no matter what T_1, T_2 and τ are, in fact
$$y^f_\infty = \lim_{t\to+\infty} y^f(t) = \lim_{s\to 0} sY^f(s) = \mu$$
- its initial slope has the same of μ if $\tau > 0$, opposite otherwise, in fact
$$\dot{y}^f(0) = \lim_{s\to+\infty} s\left(sY^f(s) - y^f(0)\right) = \lim_{s\to+\infty} sG(s) = \frac{\mu\tau}{T_1T_2}$$
- in case of cancellations, i.e. when either $\tau = T_1$ or $\tau = T_2$, the response is the one of the resulting first-order system.

When $T_1 \gg T_2$, from the responses reported in Fig. 3.7, we can further observe that

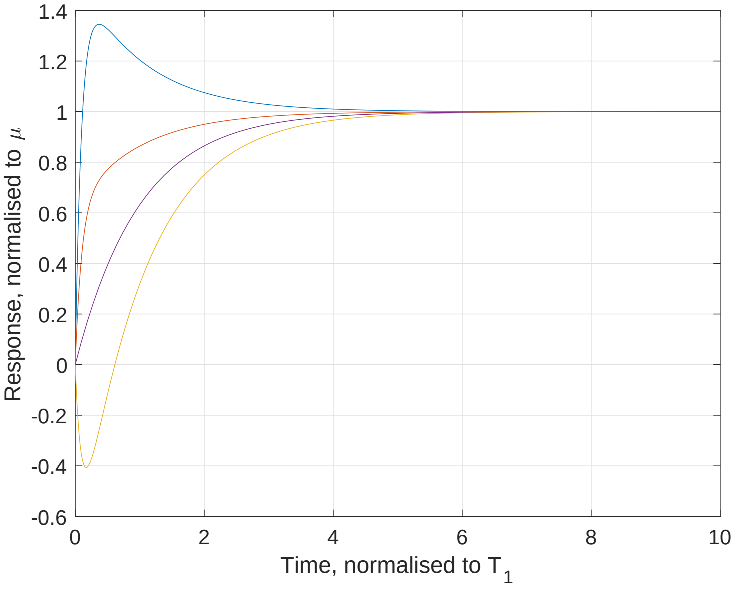


Figure 3.7: Forced step response $y^f(t)$ of a generic asymptotically stable second-order LTI SISO system with $T_2 = T_1/10$, in the cases of a dominant zero $\tau > T_1$ (blue), a zero with time constant $T_2 < \tau < T_1$ (red), the cancellation $\tau = T_2$ (purple), and zero with negative time constant $\tau < 0$ (yellow), axes are normalised to the dominant time constant T_1 and to the final value μ, respectively.

- an overshoot is present whenever $\tau > T_1 \geq T_2 > 0$;
- the negative time constant of the zero $\tau < 0$ is responsible for the undershoot;
- whenever $\varepsilon\% = 1\%$, the $\varepsilon\%$-settling time depends on T_1 only, and can be computed as $T_{s1} = \log_e(100)T_1 \approx 4.6T_1$.

Complex and conjugate poles and no zeros

We now consider the following transfer function

$$G(s) = \frac{\mu\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}, 0 < \xi < 1$$

The Laplace transform of its forced step response results as follows

$$Y^f(s) = G(s)\frac{1}{s} = \frac{\mu\omega_n^2}{s(s^2 + 2\xi\omega_n s + \omega_n^2)}$$

The corresponding analytical response results

$$y^f(t) = \mu\left(1 - \frac{1}{\sqrt{1-\xi^2}}e^{-\xi\omega_n t}\sin\left(\omega_n\sqrt{1-\xi^2}t + \arccos\xi\right)\right), t \geq 0 \quad (3.15)$$

We can observe that

- its steady-state value is $y^f_\infty = \mu$, no matter what ω_n and μ are, in fact
$$y^f_\infty = \lim_{t\to+\infty} y^f(t) = \lim_{s\to 0} sY^f(s) = \mu$$
- the initial value is null, i.e. $y^f(0) = 0$, regardless the values of μ and T, in fact
$$y^f(0) = \lim_{s\to+\infty} sY^f(s) = \lim_{s\to+\infty} G(s) = 0$$
- the initial slope is also always null, i.e. $\dot{y}^f(0) = 0$, in fact
$$\dot{y}^f(0) = \lim_{s\to+\infty} s\left(sY^f(s) - y^f(0)\right) = \lim_{s\to+\infty} sG(s) = 0$$

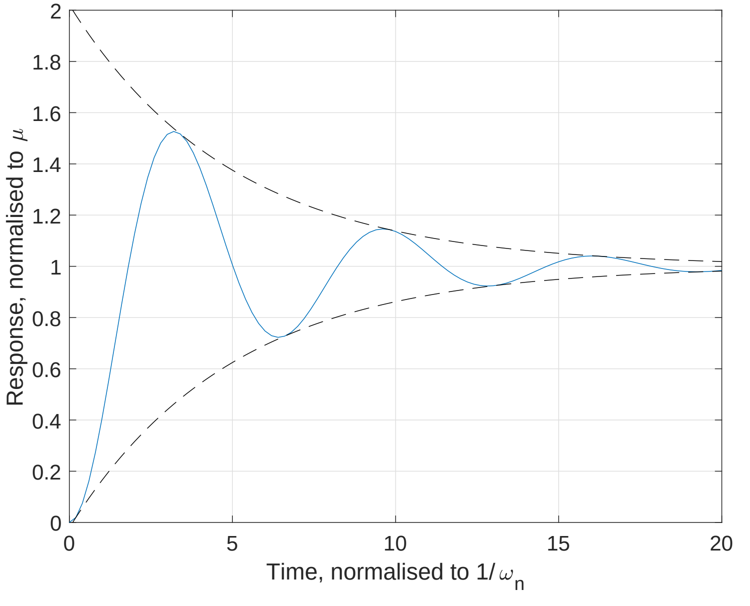


Figure 3.8: Forced step response $y^f(t)$ of a generic asymptotically stable second-order LTI SISO with complex and conjugate poles, axes are normalised to $1/\omega_n$ and to the finale value μ, respectively.

Moreover, from (3.15) we can observe that, since $|\sin(x)| \leq 1, \forall x$, we have

$$1 - \frac{1}{\sqrt{1-\xi^2}}e^{-\xi\omega_n t} \leq \frac{y^f(t)}{\mu} \leq 1 + \frac{1}{\sqrt{1-\xi^2}}e^{-\xi\omega_n t}$$

which represent lower and upper bound of the curve, respectively (see Fig. 3.8). Based on the approximated bounds, the 1%-settling time T_{s1} can be estimated imposing

$$\frac{1}{100} = \frac{1}{\sqrt{1-\xi^2}} e^{-\xi \omega_n T_{s1}}$$

which results, at least for small values of ξ (i.e. $\xi \ll 1$), is

$$T_{s1} \approx \frac{\log_e 100}{\xi \omega_n}$$

Finally, we just mention, without proving it, that the overshoot $S_{\%}$ can be computed as

$$S_{\%} = 100 e^{-\frac{\pi \xi}{\sqrt{1-\xi^2}}} \tag{3.16}$$

The graphical representation of (3.16) is reported in Fig. 3.9.

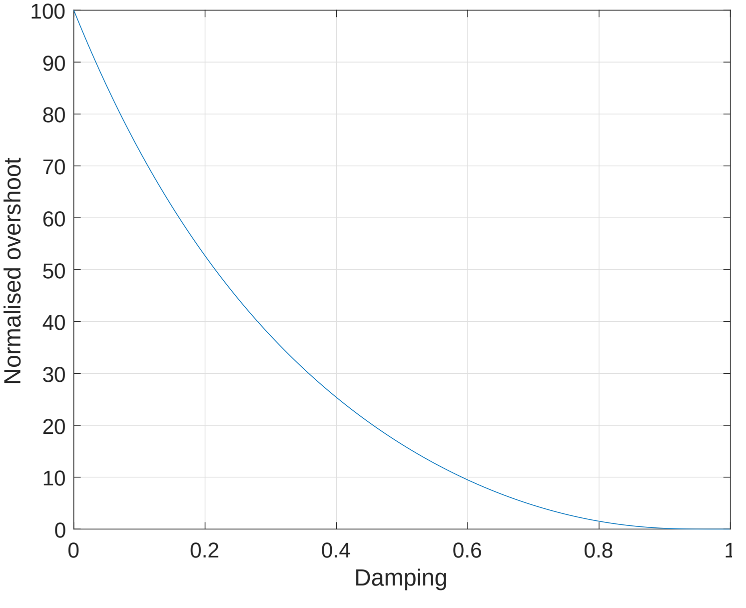


Figure 3.9: Normalised overshoot $S_{\%}$ as a function of the damping ξ.

From (3.16) and from the estimate of the 1%-settling time T_{s1}, we observe that the damping ξ plays a crucial role in both these quantities. For the same values of ω_n, the higher the damping ξ, the smaller the overshoot and the lower the settling time. Figure 3.10 reports the behaviour of (3.15) as evaluated for the same ω_n and increasing values of ξ.

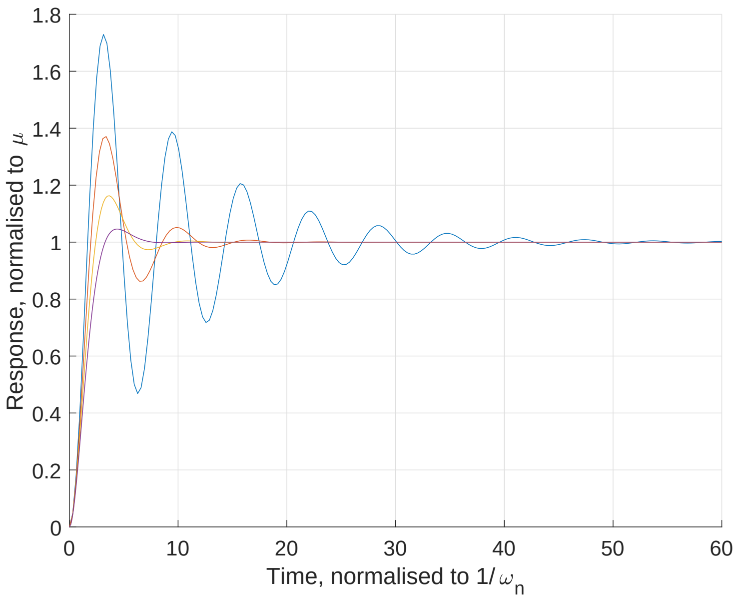


Figure 3.10: Forced step responses for the same ω_n and increasing values of ξ, axes are normalised to $1/\omega_n$ and to the finale value μ, respectively.

3.4.3 Summary

From the analyses of the forced step responses of first- and second-order asymptotically stable LTI SISO systems we can draw the following general conclusions

1. $\varepsilon\%$-settling time, and in particular 1%-settling time, is somehow proportional to either the higher time constant T in the case of real poles, or to the inverse of $\xi\omega_n$ in the case of complex and conjugate poles;

2. oscillations are present only in the case of second-order systems with complex and conjugate poles;

3. an overshoot is present either in case of complex and conjugate poles, or in the case of two real poles and one zero such that $T_2 < \tau < T_1$;

4. undershoot is only present in case of a zero with negative time constant, i.e. $\tau < 0$;

5. provided that $G(s)$ has no zeros in the origin, the steady state value y^f_∞ is equal to the generalised gain μ and the initial value is zero, i.e. $y^f(0) = 0$;

6. the initial slope $\dot{y}^f(0)$ is zero only when the relative degree is higher than one, i.e. $\nu > 1$.

Finally, the attentive reader has surely noticed that in all the discussions we assumed a null initial delay, i.e. $\tau_d = 0$. In the case a non-zero delay $\tau_d > 0$ is present in the transfer function $G(s)$ the corresponding forced step response would be just shifted ahead of τ_d seconds. The previous reasoning remains unaltered, clearly apart from the evaluation of the settling times.

Chapter 4

Frequency response of LTI SISO systems

The transfer function associated to a LTI SISO system represents a useful tool to study the forced response of the output $y^f(t)$ given an input signal $u(t)$. Recall, in fact, that given a LTI SISO system like the generic one in (3.4), if $\mathcal{L}\left[y^f(t)\right] = Y^f(s)$ and $\mathcal{L}[u(t)] = U(s)$, then

$$Y^f(s) = G(s)\,U(s)$$

which allows to easily compute the Laplace transform of the forced response of the output without relying any more on the Lagrange's formula.

4.1 Frequency response theorem

Another interesting application of the transfer function is represented by the special case of harmonic inputs, e.g.

$$u(t) = A\sin(\omega t + \psi)$$

for some values of $\omega > 0$ and ψ.

Theorem 6 (Frequency response theorem)**.** *Given a LTI asymptotically stable SISO system having transfer function $G(s)$ which is fed by input signal $u(t) = A\sin(\omega t + \psi)$ then*

$$\lim_{t\to+\infty} y^f(t) - B\sin(\omega t + \phi) = 0 \tag{4.1}$$

where $B = A\,|G(j\omega)|$ and $\phi = \psi + \angle G(j\omega)$.

The aforementioned result states that, if the input is a harmonic signal, then, after a possible transient, the forced response of the output $y^f(t)$ will be shaped as another harmonic signal **at the same frequency of the input**, with magnitude and phase shift that depend on those of the input and the transfer function $G(s)$ evaluated for $s = j\omega$, which is called **frequency response** (associated to $G(s)$).

Example 23. *[Frequency response for the RL circuit] We have previously seen that when the RL circuit in Fig. 2.3 is fed with input voltage $v(t) = \sin(\omega t)$, the trajectory of the state was the one in (2.11), i.e.*

$$x(t) = e^{-(R/L)t}x(0) + \frac{Le^{-(R/L)t} - L\cos(t) + R\sin(t)}{L^2 + R^2}, t \geq 0$$

If one considers the voltage drop on the resistor $y(t) = Rx(t)$ the corresponding trajectory can be easily computed. It is worth noticing that, for $t \to +\infty$, the trajectory of the output is such that

$$\lim_{t\to+\infty} y^f(t) - \frac{R^2\sin(t) - RL\cos(t)}{L^2 + R^2} = 0$$

which can equivalently rewritten as

$$\lim_{t\to+\infty} y^f(t) - \sqrt{\frac{R^2}{L^2 + R^2}}\sin\left(t - \mathrm{atan}\left(\frac{R}{L}\right)\right) = 0$$

Dealing with the same example, we have been able to evaluate its transfer function being the one in (3.3), i.e.

$$G(s) = \frac{R/L}{s + R/L}$$

The application of the frequency response theorem would lead to the evaluation of $G(s)$ in $s = j\omega$ with $\omega = 1$, i.e.

$$G(j1) = \frac{R/L}{j + R/L} = \frac{R/L(-j + R/L)}{1 + (R/L)^2} = \frac{R^2}{L^2 + R^2} - j\frac{RL}{L^2 + R^2}$$

which, in turn, can be rewritten in terms of its magnitude and its phase as follows

$$G(j1) = |G(j1)|\, e^{j\angle G(j1)} = \sqrt{\frac{R^2}{L^2 + R^2}}e^{-j\mathrm{atan}\left(\frac{R}{L}\right)}$$

Example 24. *[Frequency response for the mass-spring-damper system] Consider mass-spring-damper in Fig. 2.2 with $m = 1$, $k = 200$ and $r = 1$. In addition, assume that we define the elongation of the spring $x_1(t)$ as the output for the system, i.e. $y(t) = x_1(t)$. In this setting, the transfer function of the system is obtained as*

$$G(s) = \frac{1}{s^2 + s + 200}$$

Assume that one wants to find the frequency $\bar{\omega}$ such that the amplitude of the oscillation of the output is maximal when the input is a generic harmonic signal $u(t) = \sin(\omega t)$.
Since the amplitude of the output can be predicted, at least after some transient, using the frequency response theorem as $|G(j\bar{\omega})|$, we are basically looking for the solution of

$$\bar{\omega} = \mathrm{argmax}_\omega\, |G(j\omega)| = \mathrm{argmax}_\omega \left|\frac{1}{-\omega^2 + j\omega + 200}\right| = 14.1244 \approx \sqrt{\frac{k}{m}} = \sqrt{200} = 14.1421$$

Assume now that the input signal is a linear combination of several (possibly infinite) harmonics, i.e.

$$u(t) = \sum_i A_i \sin(\omega_i t + \psi_i)$$

Then, the corresponding forced response of the output $y^f(t)$ could be evaluated, once again, with the Lagrange's formula, i.e.

$$y^f(t) = \boldsymbol{C} \int_0^t e^{\boldsymbol{A}(t-\tau)} \boldsymbol{B} \sum_i A_i \sin(\omega_i \tau + \psi_i) d\tau$$

For linearity of the integral operator (i.e. for the superposition principle reported in Chapter 2), one could equivalently compute

$$y^f(t) = \sum_i \boldsymbol{C} \int_0^t e^{\boldsymbol{A}(t-\tau)} \boldsymbol{B} A_i \sin(\omega_i \tau + \psi_i)\, d\tau$$

Each component of the forced response of the output can be approximated, at least after the transient, using the frequency response theorem, yielding the following identity:

$$\lim_{t \to +\infty} y^f(t) - \sum_i A_i |G(j\omega_i)| \sin(\omega_i t + \psi_i + \angle G(j\omega_i)) = 0$$

where $G(j\omega_i)$ are multiple evaluations, for all values ω_i's, of the frequency response associated to transfer function $G(s)$.

4.2 Bode's diagrams

From the latest discussion, we have seen that typical applications of the frequency response theorem might require multiple evaluations (at different frequencies) of the frequency response of a given system.
For a given frequency $\omega \in [0, +\infty)$, the frequency response associated to transfer function $G(s)$, i.e. quantity $G(j\omega)$, is, in general a complex number with a certain magnitude and a certain phase, i.e.

$$G(j\omega) = |G(j\omega)|\, e^{j\angle G(j\omega)} = \rho(\omega)\, e^{j\theta(\omega)}$$

These two quantities, i.e. the magnitude $\rho(\omega) = |G(j\omega)|$ and the phase $\theta(\omega) = \angle G(j\omega)$, are two scalar quantities clearly depending on the argument ω. Being able to quickly draw these two functions would definitely allow for a faster evaluation of the frequency response theorem in the case of input signals with multiple harmonic components. Bode's diagrams are the graphical representation of the two functions

$$20 \log \rho(\omega) = 20 \log |G(j\omega)| \text{ and } \theta(\omega) = \angle G(j\omega)$$

as functions of ω, on logarithmic scale.

4.2.1 Bode's diagrams of the magnitude

Consider a LTI SISO system having transfer function $G(s)$ written in the form (3.8), i.e.

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1+s\tau_i)}{\prod_k (1+sT_k)} \frac{\prod_i \left(1+\frac{2\xi_{z,i}}{\omega_{z,i}}s+\frac{s^2}{\omega_{z,i}^2}\right)}{\prod_k \left(1+\frac{2\xi_{p,k}}{\omega_{p,k}}s+\frac{s^2}{\omega_{p,k}^2}\right)}$$

The magnitude of the corresponding frequency response can be computed as

$$\rho(\omega) = |G(j\omega)| = \frac{|\mu|}{|(j\omega)^g|} \frac{\prod_i |1+j\omega\tau_i|}{\prod_k |1+j\omega T_k|} \frac{\prod_i \left|1+j\frac{2\xi_{z,i}}{\omega_{z,i}}\omega+\frac{(j\omega)^2}{\omega_{z,i}^2}\right|}{\prod_k \left|1+j\frac{2\xi_{p,k}}{\omega_{p,k}}\omega+\frac{(j\omega)^2}{\omega_{p,k}^2}\right|}$$

which appears quite complicated. In turn, quantity $20\log\rho(\omega)$ can be evaluated in a much easier way, in fact

$$\begin{aligned} 20\log\rho(\omega) = {} & 20\log|\mu| - 20\log|(j\omega)^g| + \\ & \sum_i 20\log|1+j\omega\tau_i| - \sum_k 20\log|1+j\omega T_k| + \\ & \sum_i 20\log\left|1+j\frac{2\xi_{z,i}}{\omega_{z,i}}\omega+\frac{(j\omega)^2}{\omega_{z,i}^2}\right| - \sum_k 20\log\left|1+j\frac{2\xi_{p,k}}{\omega_{p,k}}\omega+\frac{(j\omega)^2}{\omega_{p,k}^2}\right| \end{aligned} \tag{4.2}$$

It should be clear that, for plotting the overall magnitude diagram, we need first to understand the plot of each possible element.

Gain

Focus first on the quantity $20\log|\mu|$. The diagram associated to this element of the sum is simply a horizontal line.

Type

The quantity $-20\log|(j\omega)^g| = -g20\log\omega$ corresponds to a line (with respect to $\log\omega$) with slope $-g20$.

Real zero

Quantity $20\log|1+j\omega\tau| = 20\log\sqrt{1+\omega^2\tau^2}$ is a bit more involved than the previous ones. Depending on ω, two asymptotic behaviours can be identified

$$20\log\sqrt{1+\omega^2\tau^2} \approx \begin{cases} 0, \omega \ll 1/|\tau|, \\ 20\log|\omega\tau|, \omega \gg 1/|\tau| \end{cases}$$

The two approximations, one valid way before the frequency of the real zero, the other way after, intersect at the frequency of the zero itself, i.e. at $\omega = 1/|\tau|$. At such a frequency, the approximation error is maximum and corresponds to $20\log\sqrt{2} = 10\log 2 \approx 3$ (see Fig. 4.1).

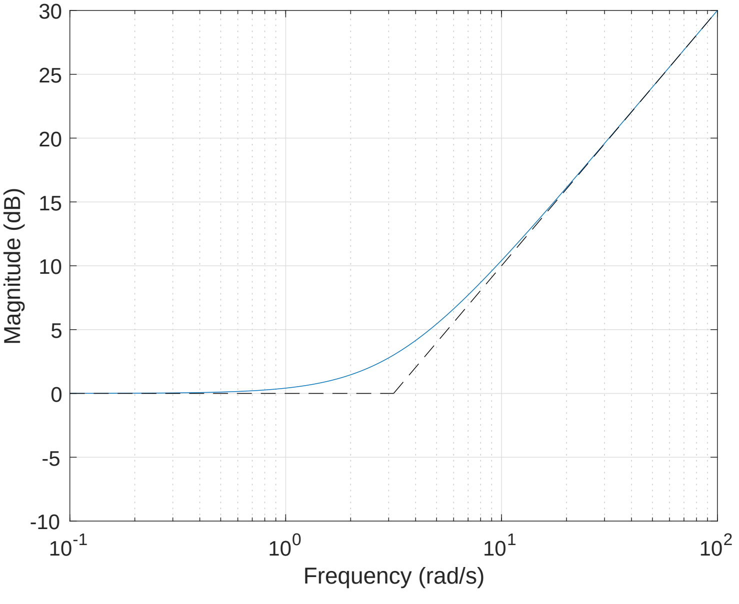


Figure 4.1: Bode's diagram of the magnitude of $1+s\tau$ with $|\tau| = 1/\sqrt{10}$.

Real pole

The plot corresponding to this element is identical to the previous one, apart from the sign (see Fig. 4.2), in fact

$$-20\log\sqrt{1+\omega^2T^2} \approx \begin{cases} 0, \omega \ll 1/|T|, \\ -20\log|\omega T|, \omega \gg 1/|T| \end{cases}$$

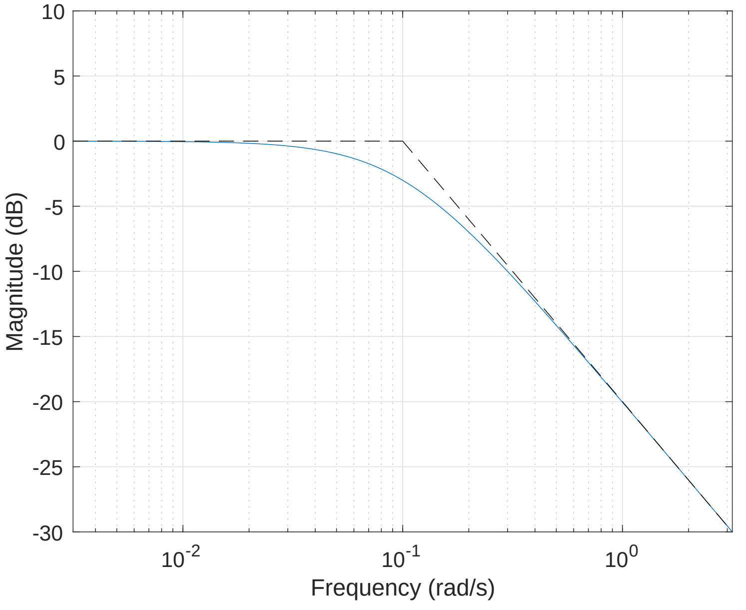


Figure 4.2: Bode's diagram of the magnitude of $(1 + sT)^{-1}$ with $|T| = 10$.

Complex and conjugate zeros

The element characterising a pair of complex and conjugate zeros

$$20 \log \left| 1 + j \frac{2\xi_z}{\omega_z}\omega + \frac{(j\omega)^2}{\omega_z^2} \right|$$

can be rearranged and approximated as follows

$$20 \log \sqrt{\left[1 - \left(\frac{\omega}{\omega_z}\right)^2\right]^2 + 4\xi_z^2 \left(\frac{\omega}{\omega_z}\right)^2} \approx \begin{cases} 0, \omega \ll \omega_z, \\ 40 \log \dfrac{\omega}{\omega_z}, \omega \gg \omega_z \end{cases}$$

In this case, the two approximations intersect corresponding to the frequency of the two complex and conjugate zeros, i.e. for $\omega = \omega_z$. The approximation error is also maximum at this frequency and corresponds to

$$20 \log \sqrt{4\xi_z^2} = 20 \log 2 |\xi_z|$$

Hence, since $|\xi_z| < 1$, the lower the damping, the higher the approximation error (see Fig. 4.3).

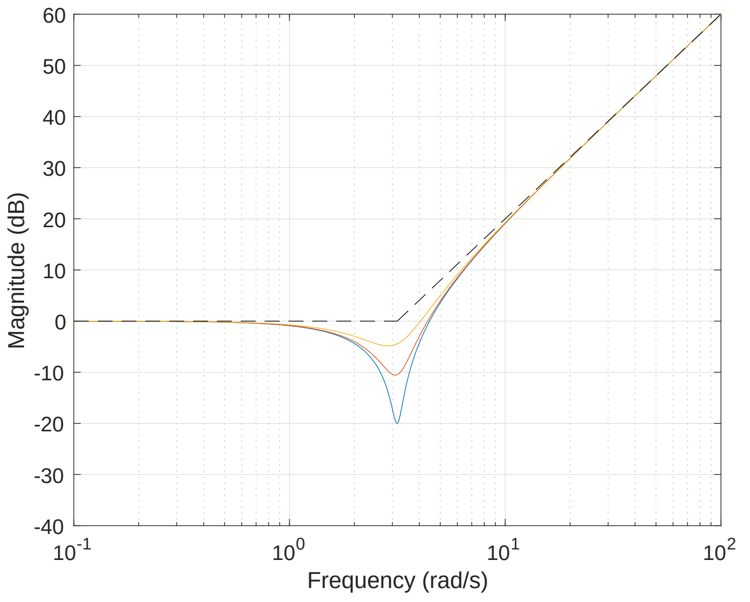


Figure 4.3: Bode's diagram of the magnitude of $1 + 2\xi_z s/\omega_z + s^2/\omega_z^2$ with $\omega_z = \sqrt{10}$ and varying $|\xi_z| \in \{0.05, 0.15, 0.30\}$.

Complex and conjugate poles

Similarly to the case of a pair of complex and conjugate zeros, the case of poles (see Fig. 4.4) results

$$-20\log\sqrt{\left[1-\left(\frac{\omega}{\omega_p}\right)^2\right]^2+4\xi_p^2\left(\frac{\omega}{\omega_p}\right)^2} \approx \begin{cases} 0, \omega \ll \omega_p, \\ -40\log\frac{\omega}{\omega_p}, \omega \gg \omega_p \end{cases}$$

Example 25. *[Bode's diagram of the magnitude] The Bode's diagram of the magnitude of*

$$G(s) = \frac{10}{s}\frac{1-s}{(1+0.1s)^2}$$

can be obtained by summing four components

$$20\log\rho(\omega) = 20\log|10| - 20\log|j\omega| + 20\log|1-j\omega| - 40\log|1+j0.1\omega|$$

The resulting magnitude diagram is shown in Fig. 4.5.

Rules for quick drawing of the magnitude diagram

From the previous example, we can observe that the resulting approximated diagram is a piecewise linear function and the slope changes corresponding to

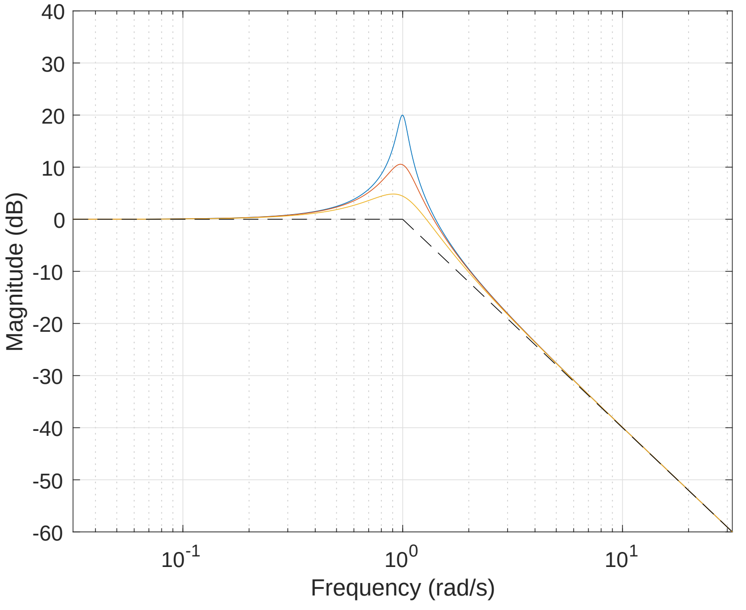


Figure 4.4: Bode's diagram of the magnitude of $\left(1 + 2\xi_p s/\omega_p + s^2/\omega_p^2\right)^{-1}$ with $\omega_p = 1$ and varying $|\xi_p| \in \{0.05, 0.15, 0.30\}$.

the frequencies of poles or zeros.
We can therefore introduce the following set of rules to quickly draw the (approximated) magnitude diagram

- at low frequencies (i.e. before the frequency of all poles and zeros, apart from those in the origin), the magnitude diagram is a line with slope $-20g$ which passes through point $(1, 20\log|\mu|)$;
- at the frequency of p real poles (zeros, respectively), the slope decreases (increases) of $20p$;
- at the frequency of p pairs of complex and conjugate poles (zeros, respectively), the slope decreases (increases) of $40p$.

Property 12. *At every frequency ω, the slope of the approximated magnitude diagram depends on the difference between zeros and poles before that frequency (accounted with their multiplicity). In particular, for $\omega \to +\infty$ the slope is the opposite of the relative degree ν of $G(s)$.*

Finally, the presence of a delay $\tau > 0$ in either the input of the output signals, does not effect the magnitude diagram, in fact

$$\left|e^{-j\omega\tau}\right| = 1, \forall\omega$$

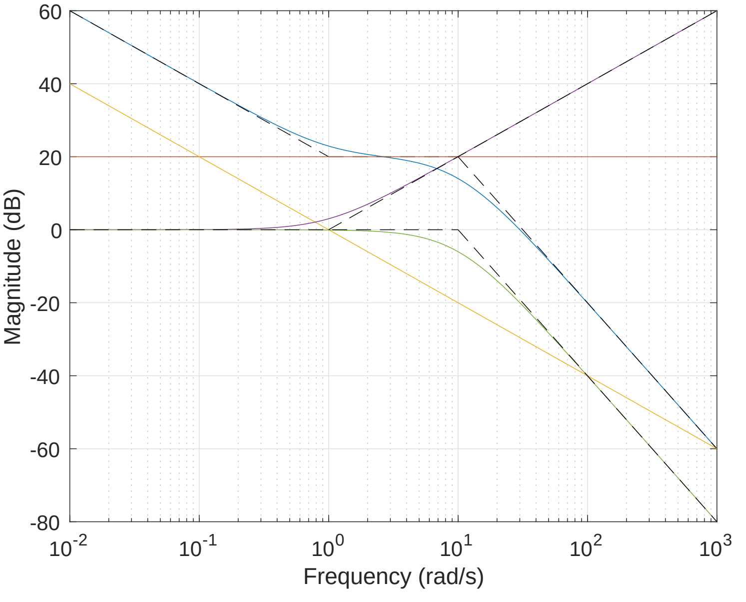


Figure 4.5: Bode's diagram of the magnitude of the frequency response of $G(s)$ (blue) and its components.

Example 26. *[Mass-spring-damper] The Bode's diagram of the magnitude of the frequency response of the mass-spring-damper system in Fig. 2.2 with $m = 1$, $k = 200$, and $r = 1$ can be obtained by observing that the two poles are complex and conjugate, in fact*

$$r^2 - 4mk = 1 - 800 < 0$$

Hence, the transfer function can be rewritten as follows

$$G(s) = \frac{1/k}{1 + 2\frac{\xi_p}{\omega_p}s + \frac{s^2}{\omega_p^2}}$$

where $\omega_p = \sqrt{k/m}$ and $\xi_p = c/(2m\omega_p)$. The corresponding Bode's diagram of the magnitude is reported in Fig. 4.6. Notice its maximum value (resonance peak) at $\omega = \bar{\omega} \approx 14.1244$ which has been computed earlier.

4.2.2 Bode's diagrams of the phase

As we have addressed how to quickly draw the (approximated) magnitude diagrams, we can introduce the corresponding reasoning to understand how to quickly sketch the (approximated) phase diagrams.

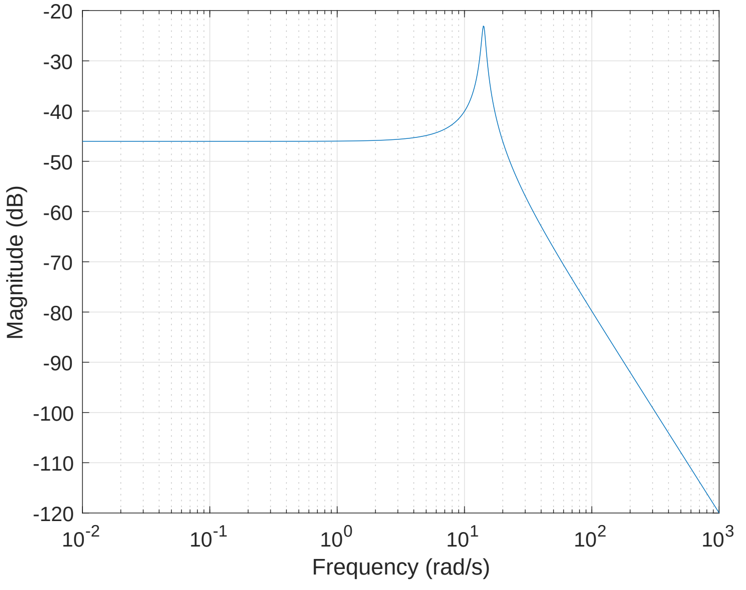


Figure 4.6: Bode's diagram of the magnitude of the mass-spring-damper system.

Consider again a LTI SISO system having transfer function $G(s)$ written in the form (3.8), i.e.

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1+s\tau_i)}{\prod_k (1+sT_k)} \frac{\prod_i \left(1 + \frac{2\xi_{z,i}}{\omega_{z,i}} s + \frac{s^2}{\omega_{z,i}^2}\right)}{\prod_k \left(1 + \frac{2\xi_{p,k}}{\omega_{p,k}} s + \frac{s^2}{\omega_{p,k}^2}\right)}$$

The phase of the corresponding frequency response can be computed as

$$\theta(\omega) = \angle G(j\omega) = \angle\mu + \angle\frac{1}{(j\omega)^g} + \sum_i \angle(1+j\omega\tau_i) + \sum_k \angle\left(\frac{1}{1+j\omega T_k}\right) + \\ + \sum_i \angle\left(1 + j\frac{2\xi_{z,i}\omega}{\omega_{z,i}} + \frac{(j\omega)^2}{\omega_{z,i}^2}\right) + \sum_k \angle\left(\frac{1}{1 + j\frac{2\xi_{p,k}\omega}{\omega_{p,k}} + \frac{(j\omega)^2}{\omega_{p,k}^2}}\right) \tag{4.3}$$

Gain

The phase of μ is simply constant (with respect to ω) and is given by:

$$\angle\mu = \begin{cases} 0, \mu > 0 \\ -\pi, \mu < 0 \end{cases}$$

Type

The phase of $1/\left(j\omega\right)^g$ is also constant and can be evaluated as

$$\angle\frac{1}{\left(j\omega\right)^g} = g\angle\left(-j\right) = -g\frac{\pi}{2}$$

Real poles and zeros

The phase contribution corresponding to real poles and zeros can be approximated as

$$\angle\left(1 + j\omega T\right)^{\pm 1} = \pm\text{atan}\left(\omega T\right) = \begin{cases} 0, \omega \ll 1/\left|T\right| \\ \pm\frac{\pi}{2}\text{sign}\left(T\right), \omega \gg 1/\left|T\right| \end{cases}$$

Notice that real poles (zeros, respectively) with a positive (negative) time constant T brings a negative (positive) phase shift. The corresponding diagrams are reported in Fig. 4.7.

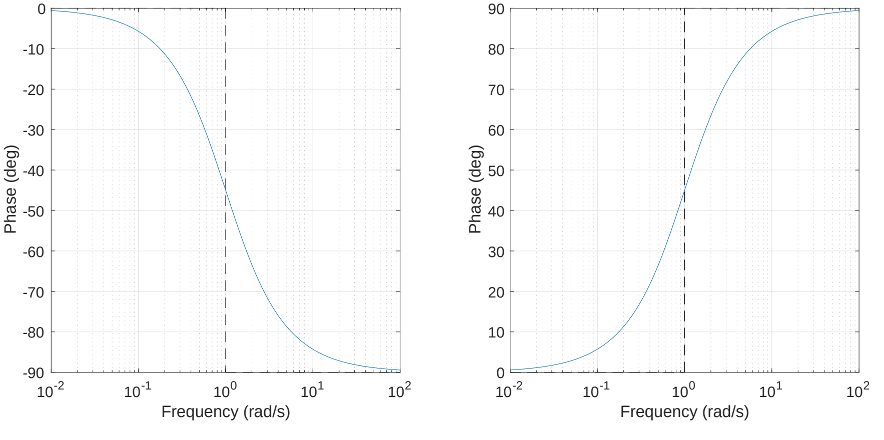


Figure 4.7: Bode's diagrams of the phase of $(1 + sT)^m$, when $m = -1, T = 1$ or equivalently $m = 1, T = -1$ (left) and $m = 1, T = 1$ or equivalently $m = -1, T = -1$ (right).

Complex and conjugate poles and zeros

Finally, the phase contribution corresponding to complex and conjugate poles and zeros can be approximated as

$$\angle\left(1+j\frac{2\xi\omega}{\omega_n}+\frac{(j\omega)^2}{\omega_n^2}\right)^{\pm1}=\begin{cases}0,\omega\ll\omega_n\\ \pm\pi\mathrm{sign}\left(\xi\right),\omega\gg\omega_n\end{cases}$$

Notice that complex and conjugate poles (zeros, respectively) with a positive (negative) damping ξ brings a negative (positive) phase shift. Beside the two asymptotic values, which do not depend on $|\xi|$, the actual shape of the phase diagram does in fact

$$\angle\left(1+j\frac{2\xi\omega}{\omega_n}+\frac{(j\omega)^2}{\omega_n^2}\right)=\angle\left(1-\frac{\omega^2}{\omega_n^2}+j\frac{2\xi\omega}{\omega_n}\right)=\mathrm{atan}\left(\frac{\frac{2\xi\omega}{\omega_n}}{1-\frac{\omega^2}{\omega_n^2}}\right)$$

Letting $x=\omega/\omega_n$, we obtain

$$\mathrm{atan}\left(2\xi\frac{x}{1-x^2}\right)$$

Though the function is not defined for $x=1$, we can notice that

$$\lim_{x\to1^-}\mathrm{atan}\left(2\xi\frac{x}{1-x^2}\right)=\lim_{x\to1^+}\mathrm{atan}\left(2\xi\frac{x}{1-x^2}\right)=\mathrm{sign}\left(\xi\right)\frac{\pi}{2}$$

hence the function can be prolonged for continuity for $x=1$. Moreover, since

$$\frac{d}{dx}\mathrm{atan}\left(2\xi\frac{x}{1-x^2}\right)=2\xi\frac{1+x^2}{\left(1-x^2\right)^2+4x^2\xi^2}$$

we finally obtain

$$\lim_{x\to1}\frac{d}{dx}\mathrm{atan}\left(2\xi\frac{x}{1-x^2}\right)=\frac{1}{\xi}$$

which clearly shows that the smaller the damping ξ, the steepest the phase diagram in the neighbourhood of ω/ω_n. The phase diagrams for complex and conjugate zeroes or poles are reported in Fig. 4.8.

Example 27. *The Bode's diagram of the phase of*

$$G\left(s\right)=\frac{10}{s}\frac{1-s}{\left(1+0.1s\right)^2}$$

can be obtained by summing four components

$$\theta\left(\omega\right)=\angle10+\angle\frac{1}{j\omega}+\angle\left(1-j\omega\right)+2\angle\left(\frac{1}{1+j0.1\omega}\right)$$

The resulting phase diagram is shown in Fig. 4.9.

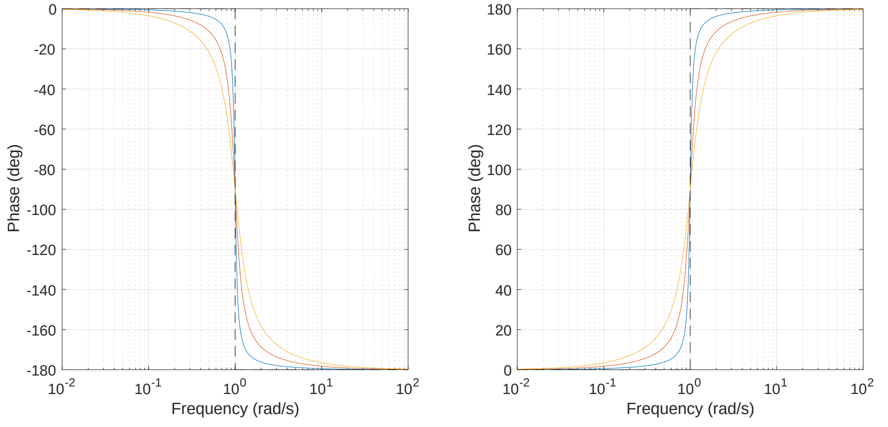


Figure 4.8: Bode's diagrams of the phase of $\left(1+2\xi_p s/\omega_p+s^2/\omega_p^2\right)^m$ with $\omega_n = 1$, when $m = -1, \xi \in \{0.05, 0.15, 0.30\}$ or equivalently $m = 1, \xi \in \{-0.30, -0.15, -0.05\}$ (left) and $m = 1, \xi \in \{0.05, 0.15, 0.30\}$ or equivalently $m = -1, \xi \in \{-0.30, -0.15, -0.05\}$ (right).

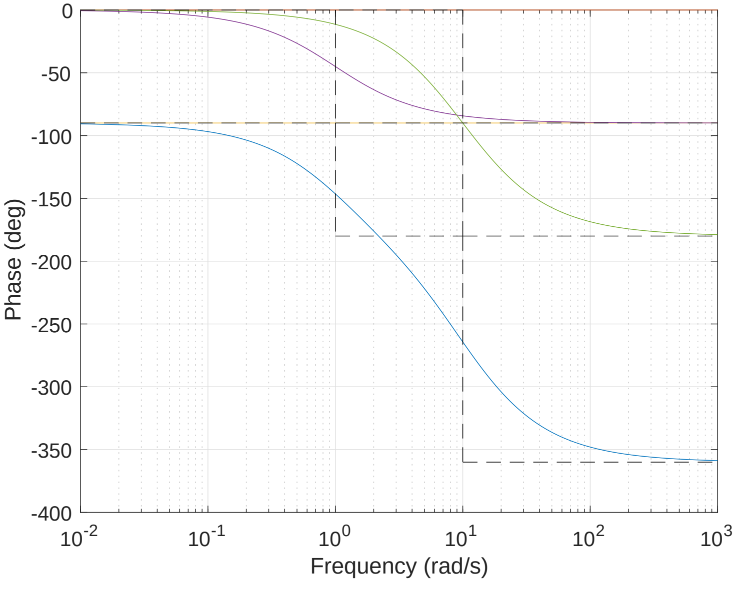


Figure 4.9: Bode's diagram of the phase of the frequency response of $G(s)$ (blue) and its components.

Rules for quick drawing of the phase diagram

From the previous example, we can observe that the resulting approximated diagram is a piecewise constant function and its value changes corresponding

to the frequencies of poles or zeros.
We can therefore introduce the following set of rules to quickly draw the (approximated) phase diagram

- at low frequencies (i.e. before the frequency of all poles and zeros, apart from those in the origin), the phase diagram is a horizontal line valued $\angle\mu - g\pi/2$;
- at the frequency of p real poles (zeros, respectively) with positive (negative) time constant, the phase steps down of $p\pi/2$;
- at the frequency of p real poles (zeros, respectively) with negative (positive) time constant, the phase steps up of $p\pi/2$;
- at the frequency of p pairs of complex and conjugate poles (zeros, respectively) with positive (negative) damping, the phase steps down of $p\pi$;
- at the frequency of p pairs of complex and conjugate poles (zeros, respectively) with negative (positive) damping, the phase steps up of $p\pi$.

Differently from the case of the magnitude diagram, the presence of a delay $\tau > 0$ in either the input or the output signals shifts the phase diagram of quantity

$$\angle e^{-\omega\tau} = -\omega\tau < 0, \forall\omega$$

which is always negative, proportional to the delay itself, and linear with respect to the frequency ω, hence exponential with respect to the $\log\omega$.

4.3 LTI SISO systems as filters

At the beginning of this Chapter, we have introduced the frequency response theorem which allows one to quickly evaluate the forced response of the output of an **asymptotically stable** LTI SISO system fed with a sinusoidal signal, at least after some transient. As mentioned, the forced response of the output will eventually be a sinusoidal signal with the same frequency $\bar{\omega}$ of the one in input and magnitude scaled by $|G(j\bar{\omega})|$, which can be immediately accessed from Bode's diagrams. We have also seen that, whenever the input signal is the superposition of several harmonics, say $\omega_i, i = 1, \ldots, \omega_N$, the forced response of the output signal can be computed by repeatedly applying the frequency response theorem to each contribution, which further justifies the role of Bode's diagram for a quick evaluation of $|G(j\omega_i)|, i = 1, \ldots, N$. It follows that, depending on their transfer function $G(s)$, and in particular depending on $|G(j\omega)|$ (i.e. the corresponding magnitude diagram), asymptotically stable LTI SISO systems act as filters, allowing certain frequencies ω to "pass" without being attenuated, i.e. if $|G(j\omega)| \approx 1$, others to "pass" after a substantial attenuation, i.e. if $|G(j\omega)| \ll 1$. Correspondingly, based on this property, some asymptotically stable LTI SISO systems can be classified as

1. **low-pass filters**, those characterised by $|G(j\omega)| \approx 1$, only for $\omega < \omega_f$, and $|G(j\omega)| \ll 1$ elsewhere;

2. **high-pass filters**, those characterised by $|G(j\omega)| \approx 1$, only for $\omega > \omega_f$, and $|G(j\omega)| \ll 1$ elsewhere;

3. **band-pass filters**, those characterised by $|G(j\omega)| \approx 1$, only for $\omega_l < \omega < \omega_h$, and $|G(j\omega)| \ll 1$ elsewhere;

4. **band-stop filters**, those characterised by $|G(j\omega)| \ll 1$, only for $\omega_l < \omega < \omega_h$, and $|G(j\omega)| \approx 1$ elsewhere.

Example 28. *[LTI SISO systems as filters] Consider the four transfer functions in the following*

$$G_1(s) = \frac{1}{1+s}, G_2(s) = \frac{s}{1+s}$$

$$G_3(s) = \frac{10s}{(1+10s)(1+0.1s)}, G_4(s) = \frac{(1+10s)(1+0.1s)}{(1+100s)(1+0.01s)}$$

The corresponding magnitude diagrams are reported in Fig. 4.10. As one can see, transfer functions $G_1(s), G_2(s), G_3(s)$, and $G_4(s)$ act as low-pass, high-pass, band-pass, and band-stop filters, respectively.

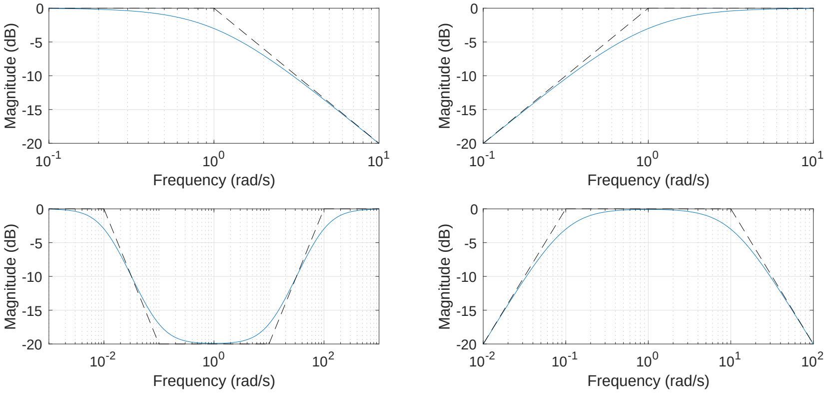


Figure 4.10: Magnitude diagrams of the frequency responses of $G_i(s)$ (from top left, in clock-wise order, $i = 1, \ldots, 4$).

For a further confirmation, consider the following input signal

$$u(t) = \sin(0.001t) + \sin(t) + \sin(1000t)$$

which feds simultaneously the four systems. Correspondingly, after some transient, one expect the following outputs

$$y_i^f(t) = \sum_{\omega_i} |G_i(j\omega_i)| \sin(\omega_i t + \angle G_i(j\omega_i)), \omega_i = 0.001, 1, 1000$$

Referring to amplitudes only, for $G_1(s)$

$$|G_1(j0.001)| \approx 1, |G_1(j1)| = \frac{\sqrt{2}}{2}, |G_1(j1000)| \approx 0$$

while for $G_2(s)$

$$|G_2(j0.001)| \approx 0, |G_2(j1)| = \frac{\sqrt{2}}{2}, |G_2(j1000)| \approx 1$$

thus confirming the low-pass and the high-pass capability of $G_2(s)$ and $G_2(s)$, respectively. Moreover, for the two remaining transfer functions $G_3(s)$ and $G_4(s)$, we have

$$|G_3(j0.001)| \approx 0, |G_3(j1)| \approx 1, |G_3(j1000)| \approx 0$$

and

$$|G_4(j0.001)| \approx 1, |G_4(j1)| = 0.1, |G_4(j1000)| \approx 1$$

which clearly confirms the behaviour of a band-pass and a band-stop filters, respectively.

Chapter 5

Block diagrams: interconnection of LTI systems

In many engineering applications, the system to be described (or controlled) is actually composed of two or more subsystems. Block diagrams are a graphical formalism to represent the interactions between these dynamical subsystems. Assuming all subsystems to be LTI SISO systems described by means of the corresponding transfer functions, this Chapter will introduce and discuss all the possible ways these systems can interact with each other. Lately, the stability characteristics of the overall system, possibly depending on the stability of each single component, will be also derived.

$$u_i(t) \longrightarrow \boxed{G_i(s)} \longrightarrow y_i^f(t)$$

Figure 5.1: Block diagram of a generic subsystem characterised by input $u_i(t)$, output $y_i(s)$, and transfer function $G_i(s)$.

5.1 Interconnected subsystems

Consider a generic LTI SISO system having input $u_i(t)$, output $y_i(t)$, and transfer function $G_i(s)$. From the previous Chapter, we know that the relationship, in terms of the Laplace transform, between the input and the forced response of the output can be written as follows

$$Y_i^f(s) = G_i(s)\, U_i(s)$$

This relationship can be graphically represented as reported in the block diagram of Fig. 5.1.

5.1.1 Subsystems in series

Assume now to have N subsystems $G_i(s), i = 1, \ldots, N$ so that the input of subsystem $i+1$ is the forced response of the output of subsystem i, i.e. $u_{i+1}(t) = y_i^f(t), \forall t$. This situation is represented (for $N = 2$) in terms of the block diagram of Fig. 5.2.

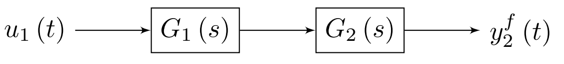


Figure 5.2: Block diagram of the series of two subsystems.

From the way the subsystems are interconnected to each other, it follows that

$$\begin{aligned} Y_N^f(s) &= G_N(s)\,U_N(s) = G_N(s)\,Y_{N-1}^f(s) = G_N(s)\,G_{N-1}(s)\,U_{N-1}(s) = \ldots \\ &= G_N(s)\,G_{N-1}(s)\ldots G_1(s)\,U_1(s) = \left(\prod_{i=N}^{1} G_i(s)\right) U_1(s) \end{aligned} \tag{5.1}$$

and therefore the transfer function of the overall system can be obtained as the **product** of the transfer functions of its components. In the example of Fig. 5.2, in fact, one would obtain

$$Y_2^f(s) = \left(\prod_{i=2}^{1} G_i(s)\right) U_1(s) = G_2(s)\,G_1(s)\,U_1(s)$$

5.1.2 Subsystems in parallel

Assume now to have N subsystems $G_i(s), i = 1, \ldots, N$ all fed with the same input, say $u(t)$ and such that the forced response of the output of the overall system $y^f(t)$ is the sum of the force response of the outputs $y_i^f(t)$ of each subsystem. This situation, for $N = 2$, is reported in the block diagram of Fig. 5.3.

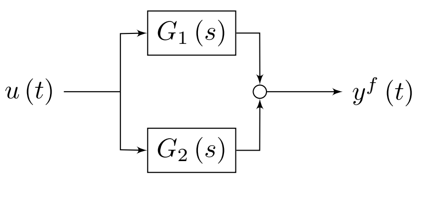


Figure 5.3: Block diagram of two subsystems in parallel.

It then follows that

$$\begin{aligned} Y^f(s) &= Y_1^f(s) + Y_2^f(s) + \cdots + Y_N^f(s) = \\ &= G_1(s)U(s) + G_2(s)U(s) + \cdots + G_N(s)U(s) = \\ &= (G_1(s) + G_2(s) + \cdots + G_N(s))U(s) = \left(\sum_{i=1}^{N} G_i(s)\right)U(s) \end{aligned} \tag{5.2}$$

and therefore the transfer function of the overall system can be obtained as the **sum** of the transfer functions of its components. In the example of Fig. 5.3, in fact, one would obtain

$$Y^f(s) = \left(\sum_{i=1}^{2} G_i(s)\right)U(s) = (G_1(s) + G_2(s))U(s)$$

5.1.3 Subsystems in feedback

Consider now exactly two subsystems, say $G_1(s)$ and $G_2(s)$, interconnected so that the output $y_1(t)$ of $G_1(s)$ is the input of $G_2(s)$, while the output $y_2(t)$ of the latter is added or subtracted to an external input $u(t)$ to determine the input of $G_1(s)$. This situation is depicted in Fig. 5.4.

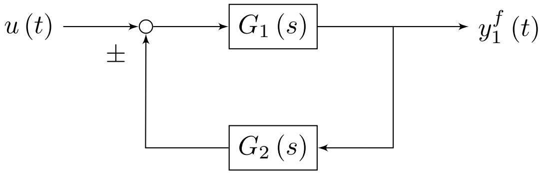


Figure 5.4: Block diagram of two subsystem in closed-loop.

The calculation of the transfer function of the overall system is slightly more involved, as compared to the previous cases. In fact

$$Y_1^f(s) = G_1(s)\left(U(s) \pm Y_2^f(s)\right) = G_1(s)\left(U(s) \pm G_2(s)Y_1^f(s)\right)$$

now entails a "circular dependency" of $Y_1^f(s)$. By isolating on the left-hand side all terms depending on $Y_1^f(s)$, we obtain

$$(1 \mp G_1(s)G_2(s))Y_1^f(s) = G_1(s)U(s)$$

and finally

$$Y_1^f(s) = \frac{G_1(s)}{1 \mp G_1(s)G_2(s)}U(s)$$

5.2 Stability analysis of interconnected subsystems

As we have seen in the previous Chapter, the stability property of LTI SISO system can be discussed, to some extent, directly from the analysis of the poles of the corresponding transfer function. However, since this method is prone to errors in the case of the lack of reachability or observability, in the following we will analyse the stability of interconnected systems referring to their state-space representation (i.e. based on eigenvalues). Therefore, we will assume each subsystem i to be represented by

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}_i(t) &= \boldsymbol{A}_i\boldsymbol{x}_i(t) + \boldsymbol{B}_i u_i(t) \\ y_i(t) &= \boldsymbol{C}_i\boldsymbol{x}_i(t) + D_i u_i(t) \end{aligned} \tag{5.3}$$

where, of course, $G_i(s) = \boldsymbol{C}_i(s\boldsymbol{I} - \boldsymbol{A}_i)^{-1}\boldsymbol{B}_i + D_i$

5.2.1 Stability of the series

For two, or more, subsystems in series, we have seen that the output of one constitutes the input of the other. Hence the equations describing the overall system can be written as follows

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}_1(t) &= \boldsymbol{A}_1\boldsymbol{x}_1(t) + \boldsymbol{B}_1 u_1(t) \\ y_1(t) &= \boldsymbol{C}\boldsymbol{x}_1(t) + D_1 u_1(t) \\ \frac{d}{dt}\boldsymbol{x}_2(t) &= \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2 u_2(t) = \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2 y_1(t) = \\ &= \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2(\boldsymbol{C}_1\boldsymbol{x}_1(t) + D_1 u_1(t)) = \\ &= \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2\boldsymbol{C}_1\boldsymbol{x}_1(t) + \boldsymbol{B}_2 D_1 u_1(t) \end{aligned} \tag{5.4}$$

and therefore, for the overall system, we have

$$\frac{d}{dt}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{0}\\ \boldsymbol{B}_2\boldsymbol{C}_1 & \boldsymbol{A}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix} + \begin{bmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}_2 D_1\end{bmatrix} u_1(t)$$

The matrix expressing the relationship between the overall state $\boldsymbol{x}(t)$ and its derivative with respect to time is clearly block-lower-triangular. Therefore, its eigenvalues are those of $\boldsymbol{A}_1$ and those of $\boldsymbol{A}_2$.

Property 13. *The series of two or more LTI SISO systems is asymptotically stable if and only if every subsystem is asymptotically stable.*

Example 29. *[Two subsystems in series] Consider the two following first-order LTI SISO subsystems*

$$\frac{d}{dt}x_1(t) = x_1(t) + u_1(t), y_1(t) = x_1(t)$$

and

$$\frac{d}{dt}x_2(t) = -x_2(t) + 2u_2(t), y_2(t) = -x_2(t) + u_2(t)$$

It is straightforward to verify that while the latter is asymptotically stable, the former is not. Moreover, they are characterised by the following transfer functions

$$G_1(s) = \frac{1}{s-1}, G_2(s) = -\frac{2}{s+1} + 1 = \frac{-2+s+1}{s+1} = \frac{s-1}{s+1}$$

While the transfer function of the overall system can be written as

$$G(s) = G_2(s) G_1(s) = \frac{1}{s-1}\frac{s-1}{s+1} = \frac{1}{s+1}$$

Despite the transfer function presents a single pole within the left half-plane, the overall system, which is of order $n = 2$, is unstable, in fact

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_1(t)$$

which is consistent with the Property.

5.2.2 Stability of the parallel

For two, or more, subsystems in parallel, the equations describing the overall system can be written as follows

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}_1(t) &= \boldsymbol{A}_1\boldsymbol{x}_1(t) + \boldsymbol{B}_1 u(t) \\ \frac{d}{dt}\boldsymbol{x}_2(t) &= \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2 u(t) \end{aligned} \tag{5.5}$$

and therefore, for the overall system, we have

$$\frac{d}{dt}\begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1(t) \\ \boldsymbol{x}_2(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{bmatrix} u(t)$$

The matrix expressing the relationship between the overall state $\boldsymbol{x}(t)$ and its derivative with respect to time is clearly block-diagonal. Therefore, its eigenvalues are those of $\boldsymbol{A}_1$ and those of $\boldsymbol{A}_2$.

Property 14. *The parallel of two or more LTI SISO systems is asymptotically stable if and only if every subsystem is asymptotically stable.*

5.2.3 Stability of feedback

In order to address the stability property of two subsystems in closed-loop consider two subsystems, which, for simplicity, are assumed to be strictly proper, i.e.

$$\begin{aligned}\frac{d}{dt}\boldsymbol{x}_1(t) &= \boldsymbol{A}_1\boldsymbol{x}_1(t) + \boldsymbol{B}_1(u(t) \pm y_2(t))\\ y_1(t) &= \boldsymbol{C}_1\boldsymbol{x}_1(t)\\ \frac{d}{dt}\boldsymbol{x}_2(t) &= \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2 y_1(t)\\ y_2(t) &= \boldsymbol{C}_2\boldsymbol{x}_2(t)\end{aligned} \tag{5.6}$$

which can be rewritten, by substituting $y_2(t)$ in the first equation and $y_1(t)$ in the second, as follows

$$\begin{aligned}\frac{d}{dt}\boldsymbol{x}_1(t) &= \boldsymbol{A}_1\boldsymbol{x}_1(t) + \boldsymbol{B}_1(u(t) \pm \boldsymbol{C}_2\boldsymbol{x}_2(t))\\ \frac{d}{dt}\boldsymbol{x}_2(t) &= \boldsymbol{A}_2\boldsymbol{x}_2(t) + \boldsymbol{B}_2\boldsymbol{C}_1\boldsymbol{x}_1(t)\end{aligned} \tag{5.7}$$

Therefore, the overall system is characterised by

$$\frac{d}{dt}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}_1 & \pm\boldsymbol{B}_1\boldsymbol{C}_2\\ \boldsymbol{B}_2\boldsymbol{C}_1 & \boldsymbol{A}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1(t)\\ \boldsymbol{x}_2(t)\end{bmatrix} + \begin{bmatrix}\boldsymbol{B}_1\\ \boldsymbol{0}\end{bmatrix}u(t)$$

Differently from the two previous cases of the series and the parallel, the matrix expressing the relationship between the overall state $\boldsymbol{x}(t)$ and its derivative with respect to time is neither block-lower-triangular nor block-diagonal. Hence, the stability of the closed-loop system is anything but straightforward.

Example 30. *[Two subsystems in closed-loop] Consider the two following asymptotically stable LTI SISO systems*

$$\frac{d}{dt}x_1(t) = -x_1(t) + 2u_1(t), y_1(t) = 3x_1(t) - 3u_1(t)$$

and

$$\frac{d}{dt}x_2(t) = -x_2(t) + u_2(t), y_2(t) = x_2(t)$$

which correspond to the following transfer functions

$$G_1(s) = 3\frac{1-s}{1+s}, G_2(s) = \frac{1}{1+s}$$

Once arranged in negative feedback, the dynamics of the overall system can be written as follows

$$\begin{aligned}\frac{d}{dt}x_1(t) &= -x_1(t) + 2(u(t) - y_2(t)) = -x_1(t) - 2x_2(t) + 2u(t)\\ \frac{d}{dt}x_2(t) &= -x_2(t) + y_1(t) = -x_2(t) + 3x_1(t) - 3u_1(t) =\\ &= 3x_1(t) - x_2(t) - 3(u(t) - y_2(t)) = 3x_1(t) + 2x_2(t) - 3u(t)\end{aligned} \tag{5.8}$$

which can be rewritten as

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 2 \\ -3 \end{bmatrix} u(t)$$

The corresponding characteristic equation is the following one

$$0 = \det(s\boldsymbol{I} - \boldsymbol{A}) = (s+1)(s-2) - 6 = s^2 - s + 4$$

which is solved by

$$s_{1,2} = \frac{1}{2} \pm j\frac{\sqrt{15}}{2}$$

clearly showing that the overall system is unstable (despite the asymptotic stability of the two subsystems).

In turn, consider

$$\frac{d}{dt}x_1(t) = x_1(t) + u_1(t), y_1(t) = x_1(t), G_1(s) = \frac{1}{1-s}$$

which is clearly unstable, and $y_2(t) = 3u_2(t) = 3y_1(t) = 3x_1(t)$ then the overall system can be written as

$$\frac{d}{dt}x_1(t) = x_1(t) + (u(t) - y_2(t)) = -2x_1(t) + u(t)$$

which is, in turn, asymptotically stable (despite the instability of $G_1(s)$).

From the previous example, it should be clear that the analysis of the stability property of feedback systems is actually more involved than the one for the series or the parallel systems. In addition, we have also seen that two asymptotically stable subsystems interconnected in negative feedback can produce an unstable system. In turn, an unstable subsystem can generate a closed-loop system which is asymptotically stable. A more formal analysis of the stability property of closed-loop systems with negative feedback will be given in the next Chapter.

Part II

Feedback control systems

Chapter 6

LTI SISO control systems

Consider a dynamical system consisting in a mass sliding with friction on a horizontal surface (Fig. 6.1). The mass is actuated by force $f(t)$ and its velocity $v(t)$ is accessible via a proper sensor. Moreover, an exogenous, yet non-measurable, disturbance force $w(t)$ might also act on the mass.

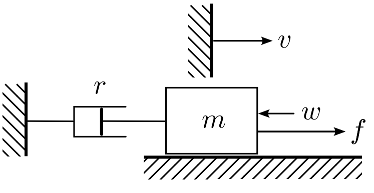


Figure 6.1: Controlled velocity of a mass with friction and disturbance.

The equation governing the behaviour of the system can be written as

$$\frac{d}{dt}v(t) = -\frac{r}{m}v(t) + \frac{1}{m}(f(t) - w(t)) \tag{6.1}$$

The transfer function from the net driving force $f(t) - w(t)$ and the velocity $v(t)$ is the following one

$$G(s) = \frac{1}{ms + r}$$

Similarly to the motivating problem in Chapter 1, assume one wants to select the actuating force $f(t)$ to let the mass travel at a desired constant velocity $\bar{v}$. A simple solution would be to set

$$f(t) = \hat{r}\bar{v}, \forall t \geq 0 \tag{6.2}$$

where $\hat{r}$ represents the best estimate of r. Further assume that the mass is already travelling at speed $\bar{v}$, i.e. $v(0) = \bar{v}$, and a disturbance force of amplitude $\bar{w}$ starts being applied to the system at time $t = 0$, i.e. $w(t) = \bar{w}, t \geq 0$. Using the Lagrange's formula, one can easily compute the corresponding response of the velocity $v(t)$ as

$$v(t) = e^{-(r/m)t}\bar{v} + \frac{\hat{r}\bar{v} - \bar{w}}{r}\left(1 - e^{-(r/m)t}\right), t \geq 0$$

which is characterised, at steady-state, by

$$\lim_{t \to +\infty} v(t) = \frac{\hat{r}\bar{v} - \bar{w}}{r}$$

Observe that

- the steady-state velocity will be influenced by both the uncertainty on the parameter r and by the amplitude of the external disturbance $\bar{w}$;
- in absence of the disturbance $w(t)$, the mass will reach the desired velocity only in case of perfect knowledge of r, i.e. $\hat{r} = r$;
- the transient of the velocity will depend on the value r/m, which is a combination of the parameters of the mechanical system.

The careful reader should have noticed that the reported properties are exactly the ones of the motivating example in Chapter 1, which have now been discovered using the tools of system theory, rather than relying on an empirical basis.

Consider now a simple algorithm that takes the measured value of $v(t)$ and actuates a force $f(t)$ that is proportional to the difference between the desired value $\bar{v}$ and the actual (measured) value $v(t)$, i.e.

$$f(t) = k(\bar{v} - v(t)) \tag{6.3}$$

where $k > 0$ is a design parameter. Note that (6.3) might represent the alternative cruise control strategy proposed by the second team in the motivating example of Chapter 1. By substituting (6.3) in (6.1), one obtains

$$\frac{d}{dt}v(t) = -\frac{k+r}{m}v(t) + \frac{1}{m}(k\bar{v} - w(t)) \tag{6.4}$$

Repeating the previous calculations, but now referring to (6.4), one would obtain

$$v(t) = e^{-\frac{k+r}{m}t}\bar{v} + \frac{k\bar{v} - \bar{w}}{k+r}\left(1 - e^{-\frac{k+r}{m}t}\right), t \geq 0$$

and

$$\lim_{t \to +\infty} v(t) = \frac{k\bar{v} - \bar{w}}{k+r}$$

which can be made arbitrarily close to $\bar{v}$ by increasing k, regardless the value of r and the presence of the disturbance $w(t)$. Moreover, the transient will now depend on $(k+r)/m$, which is now a tunable parameter.

6.1 The role of feedback

Still referring to the example of Fig. 6.1 and to the corresponding equation (6.4), we have noticed that by introducing an explicit dependency between the actuated input of the system $f(t)$ and its output (state) $v(t)$ several benefits could be obtained, namely

1. the **duration of transients** could be made arbitrarily fast by increasing k;
2. the effect of **parametric uncertainties** (related to the impossibility to have an exact estimate of r) could be arbitrarily reduced by increasing k; and
3. the **effect of the disturbance** $w(t)$ on $v(t)$ could be made arbitrarily small, again by increasing k.

These features were not achieved by the first strategy proposed in (6.2).

The two proposed strategies to achieve the goal of making the velocity of the mass $v(t)$ as similar as possible to the desired value $\bar{v}$ can be represented by means of the block diagrams in Figs. 6.3 and 6.2, respectively.

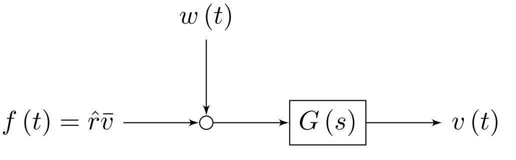


Figure 6.2: Block diagram of (6.1) together with (6.2).

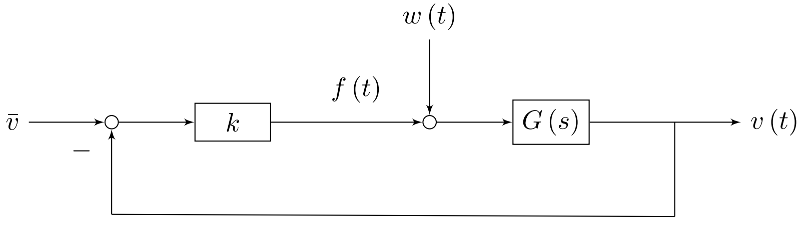


Figure 6.3: Block diagram of (6.4).

The former (Fig. 6.2), that we will call **open-loop strategy** was severely influenced by uncertainties and disturbances, while the latter (Fig. 6.3), **closed-loop strategy** or **feedback control**, was able to mitigate the effects of uncertainties and disturbances as well as capable of changing the duration of transients. The reason behind the names **closed-loop** and **feedback** simply stands in the presence of a "circular dependency" between variables. If in both cases the output variable $v(t)$ depends on $f(t)$, in the closed-loop strategy the actuated input $f(t)$ depends itself on the measured output $v(t)$ (see again Fig. 6.3).

The role of feedback was independently discovered or experienced in different fields of engineering among the centuries. Modern studies are usually assumed to be initiated by Harold Stephen Black who adopted feedback to reduce the distortion in amplifiers adopted in telephonic transmissions (1927, circa). James Clerk Maxwell, in his famous paper "On governors" (1868), proved already the possibility to mathematically interpret the capability of some machines to keep their spinning velocity approximately constant.
On the other hand, the role of feedback in changing the response of a system was already exploited few centuries before. Considering a mass M subject to gravity, one can simply write

$$\frac{d}{dt}v(t) = \frac{1}{M}(Mg + f(t))$$

where $v(t)$ is the falling velocity, g is the free-fall acceleration and $f(t)$ comprises all forces other than gravity. A parachute, an invention dating back probably to the 1400's, is a device creating an increased drag force opposed to the free-falling velocity, i.e. $f(t) = -rv(t), R > 0$.

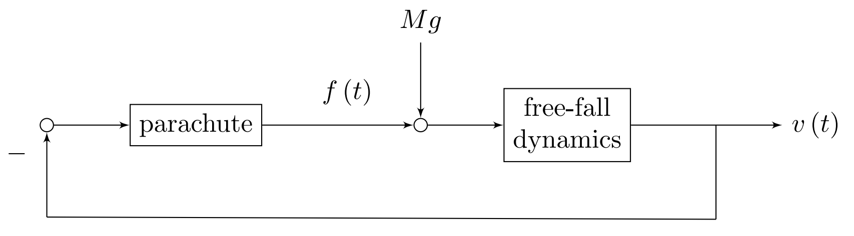


Figure 6.4: Block diagram showing how a parachute can be interpreted as a feedback system.

6.2 Stability of LTI SISO control systems

Consider the block diagram reported in Fig. 6.5 as representative of a generic feedback control system, where the series of two LTI SISO systems represented by their transfer functions $R(s)$ and $G(s)$, respectively, is arranged in a closed-loop system with negative feedback.

Define the **loop transfer function** $L(s)$ as the product of all transfer functions appearing within the loop, i.e. $L(s) = R(s)G(s)$.
Assume that the initial state of all the subsystems (i.e. both R and G) is zero.

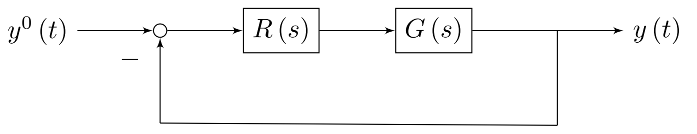


Figure 6.5: Block diagram of a simple closed-loop system.

In Chapter 3, we have seen that given the input $y^0(t)$ and its Laplace transform $Y^0(s) = \mathcal{L}\left[y^0(t)\right]$, the forced response of the output $y^f(t)$ can be computed by resorting to the transfer function of the (closed-loop) system, i.e.

$$y^f(t) = \mathcal{L}^{-1}\left[Y^f(s)\right] = \mathcal{L}^{-1}\left[\frac{L(s)}{1+L(s)}Y^0(s)\right]$$

A natural question arises: is this trajectory asymptotically stable? That is: if the initial state of the system is perturbed from zero, would the corresponding trajectory still look like $y^f(t)$, at least asymptotically?

To answer this question, one has to study the stability of the closed-loop system which is characterised by transfer function $L(s)/(1+L(s))$, as we have seen in the previous Chapter.

Example 31. *[Closed-loop stability with the Routh-Hurwitz criterion] With reference to the block diagram of Fig. 6.5, consider the following transfer function*

$$G(s) = \frac{1}{(1+s)^3}$$

and $R(s) = k$. The corresponding loop transfer function is

$$L(s) = R(s)G(s) = \frac{k}{(1+s)^3}$$

which is asymptotically stable. In turn, closed-loop transfer function is

$$\frac{L(s)}{1+L(s)} = \frac{\frac{k}{(1+s)^3}}{1+\frac{k}{(1+s)^3}} = \frac{k}{s^3+3s^2+3s+1+k}$$

Applying the Routh-Hurwitz criterion to its denominator, we obtain the following table

$$\begin{array}{ccc}
1 & 3 & 0 \\
3 & 1+k & 0 \\
-\dfrac{1+k-9}{3} & 0 & \\
1+k & &
\end{array}$$

Therefore for the closed-loop system to be asymptotically stable, a necessary and sufficient condition is that

$$\begin{cases} 1+k-9<0 \\ 1+k>0 \end{cases}$$

As a conclusion, despite the asymptotic stability of the loop transfer function $L(s)$, the closed-loop is asymptotically stable for $-1<k<8$, only.

6.2.1 Role of cancellations in closed-loop stability

Assuming transfer functions $R(s)$ and $G(s)$ being of order n_R and n_G respectively. In the calculation of $L(s)$ it might happen that a term in the numerator of $R(s)$ simplifies with a term in the denominator of $G(s)$. Under this condition, the order of the loop transfer n_L function would be $n_L < n_R + n_G$.

In the next example, we further investigate the role of cancellations happening in the right half-plane between the transfer function of the controller $R(s)$ and the one of the system $G(s)$.

Example 32. *[Cancellations within the loop transfer function] Consider a system G characterised by the following equations*

$$\begin{aligned} \frac{d}{dt}x_1(t) &= x_1(t) + u(t) \\ y(t) &= -x_1(t) \end{aligned} \tag{6.5}$$

and a controller R

$$\begin{aligned} \frac{d}{dt}x_2(t) &= y^0(t) - y(t) \\ u(t) &= x_2(t) + y(t) - y^0(t) \end{aligned} \tag{6.6}$$

The transfer functions associated to the two subsystems are the following ones

$$R(s) = \frac{1-s}{s}, G(s) = \frac{1}{1-s}$$

and are both first-order transfer functions, i.e. $n_R = n_G = 1$. The loop transfer function

$$L(s) = R(s)\,G(s) = \frac{1}{s}$$

also results in a first-order transfer function, i.e. $n_L < n_R + n_G$. In turn, the closed-loop system is represented by the following (first-order) transfer function

$$\frac{L(s)}{1+L(s)} = \frac{\frac{1}{s}}{1+\frac{1}{s}} = \frac{1}{1+s}$$

On the other hand, combining the equations of G and R clearly results in the following second-order system:

$$\begin{aligned} &\frac{d}{dt}x_1(t) = x_1(t) + u(t) = x_1(t) + x_2(t) + y(t) - y^0(t) = x_2(t) - y^0(t) \\ &\frac{d}{dt}x_2(t) = y^0(t) - y(t) = y^0(t) + x_1(t) \\ &\quad y(t) = -x_1(t) \end{aligned} \tag{6.7}$$

Considering variable $y^0(t)$ as input, and $y(t)$ as output, the matrices associated to the closed-loop system are the following ones

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} -1 & 0 \end{bmatrix}$$

*Notice, first of all, that the closed-loop system is **unstable**. The characteristic polynomial of* $\boldsymbol{A}$ *is in fact*

$$\det(\lambda \boldsymbol{I} - A) = (\lambda - 1)(\lambda + 1)$$

which clearly shows the eigenvalue $\lambda = 1$ *(which is not a pole of the closed-loop system). The closed-loop system is indeed non-reachable, in fact*

$$\boldsymbol{K_R} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{AB} \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

which is not full-rank.
To further confirm this finding, consider the following change of variables

$$\boldsymbol{z}(t) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \boldsymbol{x}(t)$$

and the corresponding equations

$$\begin{aligned} \frac{d}{dt} z_1(t) &= z_1(t) \\ \frac{d}{dt} z_2(t) &= -z_2(t) + 2y^0(t) \end{aligned} \tag{6.8}$$

The new representation of the closed-loop system clearly highlights the lack of reachability of state variable $z_1(t)$ *from the input* $y^0(t)$*, as well as its instability.*

In the previous example, we have analysed the effects of cancellation between terms in the numerator of $R(s)$ and corresponding terms in the denominator of $G(s)$. Similar conclusion could have been found in the opposite scenario. We have seen that cancellations are making part of the open-loop system either non-reachable or non-observable (or both). The reason why the closed-loop system was unstable, even though it was not visible from the closed-loop transfer function, can be easily justified resorting to the Kalman's decomposition discussed in Chapter 2.

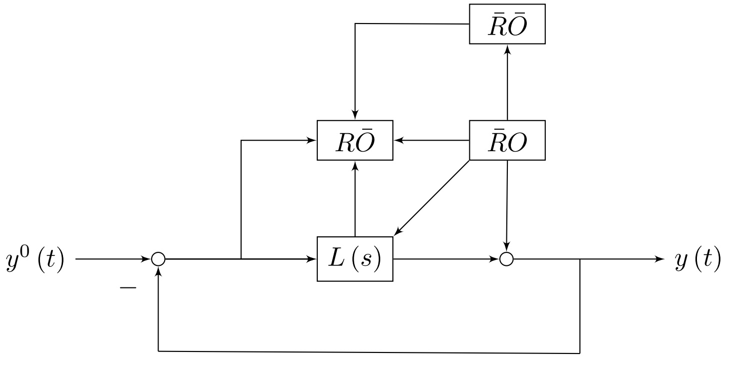


Figure 6.6: Block diagram of the closed-loop system, highlighting the Kalman's decomposition.

Figure 6.6 reports the decomposition of the closed-loop system. Transfer function $L(s)$ corresponds to the part which is simultaneously reachable and observable. Possible cancellations between $R(s)$ and $G(s)$ are no longer part of $L(s)$ and, for this reason, are outside the feedback loop. As we have seen, feedback is capable of altering the stability of a system. On the other hand, if parts of the system are simultaneously unstable and outside the control loop, they will remain as such.

6.3 Closed-loop stability analysis with Bode's diagrams

Focus again on the closed-loop system of Fig. 6.5. We have already seen that a necessary condition for its stability is that no cancellations between $R(s)$ and $G(s)$ occur within the left half-plane. Under this condition, one may refer to the Routh-Hurwitz criterion to analyse the stability of such system. As a matter of fact, the Routh-Hurwitz criterion is still applicable.

However, as we will see in the following example, we better look for something easier to be applied.

Example 33. *[Parametric stability analysis of a closed-loop system] Consider the system*

$$G(s) = \frac{1}{(1+s)^2}$$

which has to be controlled in feedback using

$$R(s) = \frac{\alpha}{s+\beta}$$

where α, β are parameters to be tuned. The closed-loop characteristic equation results the following one

$$s^3 + (2+\beta)s^2 + (1+2\beta)s + \alpha + \beta = 0 \tag{6.9}$$

while the corresponding Routh-Hurwitz table is

$$\begin{array}{ccc} 1 & (1+2\beta) & 0 \\ 2+\beta & \alpha+\beta & 0 \\ -\dfrac{\alpha+\beta-(1+2\beta)(2+\beta)}{2+\beta} & 0 & \\ \alpha+\beta & & \end{array}$$

and the corresponding region in the $\alpha\beta$-plane corresponding to closed-loop asymptotic stability is reported in Fig. 6.7.

If this strategy seems already quite involved, imagine using this tool to assume small variations on the parameters characterising the dynamics of the system to be controlled, i.e.

$$G(s) = \frac{1+\varepsilon}{(1+s)^2}$$

The example just reported clearly indicates that the Routh-Hurwitz criterion, though being applicable, is not suitable to be used as a guideline to select the parameters of the controller $R(s)$. Moreover, still in the example,

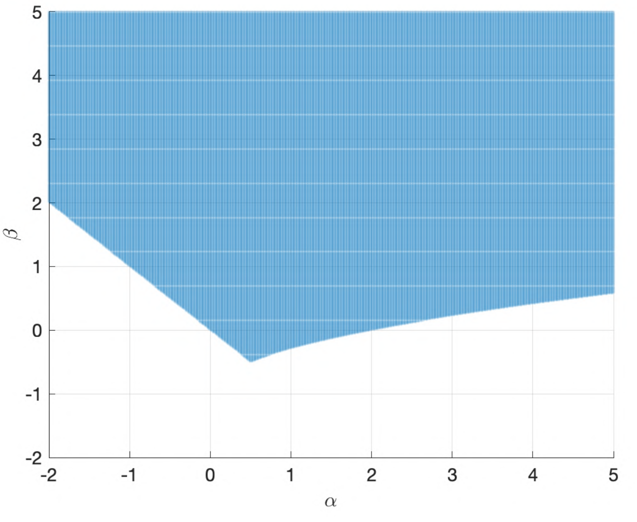


Figure 6.7: Region of the $\alpha\beta$-plane for which the closed-loop system characterised by characteristic equation in (6.9) results asymptotically stable.

the Routh-Hurwitz criterion was applicable since at least the order n_R of the transfer function characterising the controller was given. How would it be possible to apply this method if no hypothesis on the structure of the controller has been made?
As we will see next, the graphical representations of a frequency response (i.e. Bode's diagrams, see Chapter 4) come to help.

In the previous examples, we have seen that the stability of the closed-loop system depends on the roots of the (closed-loop) characteristic polynomial, i.e. the solutions of the equation

$$1 + L(s) = 0$$

which can be clearly rewritten as

$$L(s) = -1$$

Hence, how the loop-transfer function $L(s)$ is related to the value -1 (which can be also alternatively written as $e^{-j\pi}$, i.e. the complex number with unit magnitude and phase $-\pi$) represents the key point in addressing the stability property of a closed-loop system.

6.3.1 Phase margin criterion

Consider the block diagram of Fig. 6.8 where, for brevity, $L(s) = R(s)\,G(s)$ and assume no cancellations within the right half-plane between $R(s)$ and $G(s)$. The following result will show that the stability of the **closed-loop system** can be discussed relying only on the graphical representations of the frequency response of $L(s)$, i.e. its Bode's diagrams.

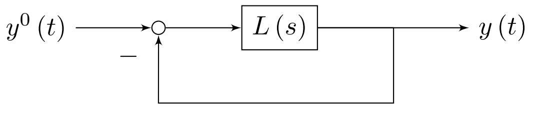


Figure 6.8: Block diagram of a simple closed-loop system, with $L(s) = R(s)\,G(s)$.

Theorem 7 (Bode's stability criterion)**.** *Given that $L(s) = R(s)\,G(s)$ has no cancellations within the closed right half-plane and **no poles within the open right half-plane**, and assume*

1. *the generalised gain μ_L of $L(s)$ is positive, i.e. $\lim_{s\to 0} s^{g_L} L(s) > 0$;*
2. *the magnitude diagram of $L(j\omega)$ has **exactly one intersection** with the 0 dB axis from above to below occurring at frequency ω_c (**crossover frequency**), i.e. $\exists! \omega_c : |L(j\omega_c)| = 1,\ |L(j\omega)| > 1, \forall \omega < \omega_c$ and $|L(j\omega)| < 0, \forall \omega > \omega_c$.*

*define ψ_m **phase margin** the quantity*

$$\psi_m = \pi - |\angle L(j\omega_c)|$$

then the closed-loop system of Fig. 6.8 is asymptotically stable if and only if

$$\psi_m > 0$$

Example 34. *We have already seen that the closed-loop system characterised by $R(s) = k$ and*

$$G(s) = \frac{1}{(1+s)^3}$$

is asymptotically stable for $-1 < k < 8$. Consider then $k = 5$ and correspondingly

$$L(s) = \frac{5}{(1+s)^3}$$

Bode's diagrams of $L(j\omega)$ are reported in Fig. 6.9 from which one can clearly notice that all the hypotheses for the applicability of the Bode's criterion are satisfied with $\omega_c \approx 1.4$ rad/s and $\psi_m = \pi - |\angle L(j\omega_c)| = 17\,\mathrm{deg} > 0$ from which the asymptotic stability of the closed-loop system directly follows.

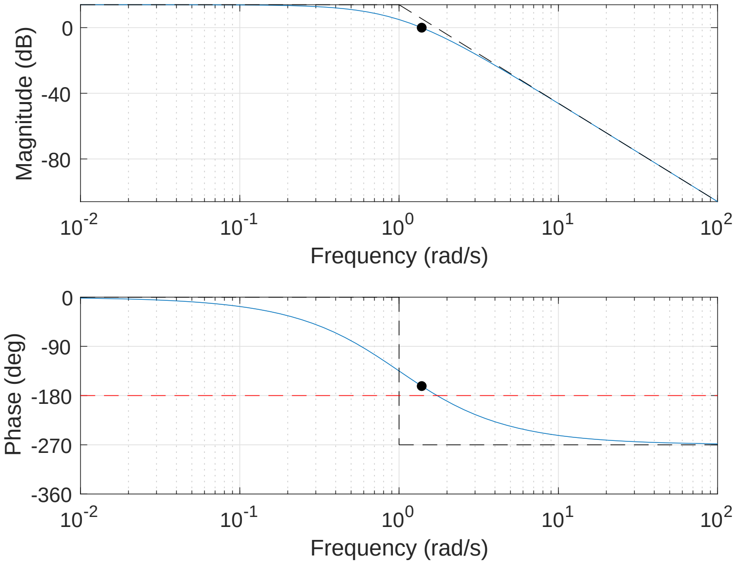


Figure 6.9: Bode's diagrams of $L(j\omega)$, from which $\omega_c \approx 1.4\ rad/s$ and $\psi_m = 17\,\mathrm{deg}$.

Example 35. *[Closed-loop stability analysis with the Bode's criterion] Consider, as a further example, the following pair*

$$G(s) = \frac{1}{1+s}, R(s) = 1000\frac{1+0.1s}{s}$$

From the Bode's diagrams in Fig. 6.10 one can notice that the closed-loop system is expected to be asymptotically stable with $\omega_c \approx 100\ rad/s$ and $\psi_m \approx 85$ deg.
Consider now a slightly modified transfer function $G(s)$, i.e.

$$\hat{G}(s) = \frac{1+\varepsilon}{1+s}, 0 < \varepsilon \ll 1$$

The Bode's diagram of the phase of $\hat{L}(j\omega)$ will remain unmodified since

$$\angle\hat{L}(j\omega) = \angle(1+\varepsilon)L(j\omega) = \angle L(j\omega)$$

In turn, the Bode's diagram of the magnitude will be slightly shifted vertically. Moreover, notice that

$$\lim_{s\to+\infty} \hat{L}(s) - \frac{100(1+\varepsilon)}{s} = 0$$

hence in the case of the slightly modified system $\hat{G}(s)$, the corresponding crossover would result approximately the same

$$\omega_c : \left|\hat{L}(j\omega_c)\right| = 1 \Rightarrow \omega_c \approx 100(1+\varepsilon)$$

and similarly for the phase margin.

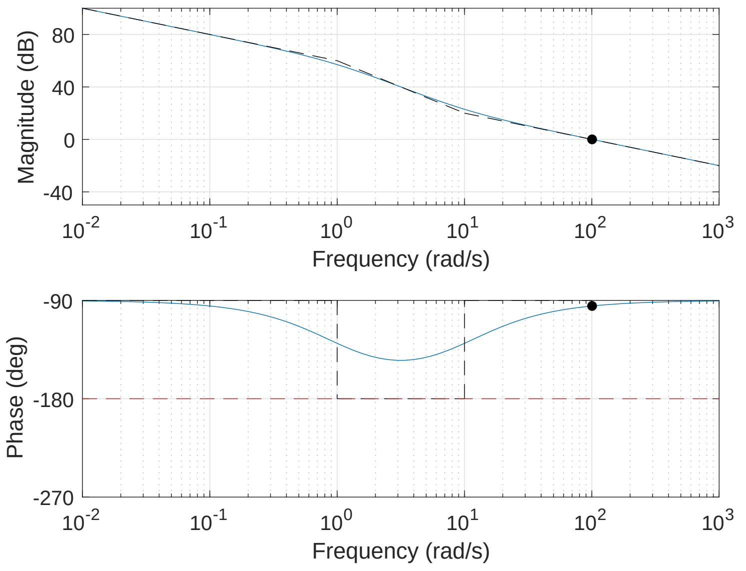


Figure 6.10: Bode's diagrams of $L(j\omega)$, from which $\omega_c \approx 100\ rad/s$ and $\psi_m = 85\,\text{deg}$.

In the previous example, we have seen that the Bode's stability criterion can be adopted to check the asymptotic stability of a closed-loop system by graphical inspection of the Bode's diagram of $L(j\omega)$. In addition, the introduction of a parametric uncertainty, and the reasoning on its influence on the Bode's diagram of $L(j\omega)$, appeared to have little influence on the stability of the closed-loop system.

As a conclusion, we can assert that the Bode's criterion, rather than the Routh-Hurwitz criterion (which however had no limits in its applicability), seems to be more appropriate for the analysis of the stability property of a closed-loop system (despite its additional hypotheses on the frequency response of $L(s)$).

Finally, notice that whenever $\psi_m = 0$ one would have, at frequency ω_c, the following phase

$$\angle L(j\omega_c) = -\pi$$

and therefore, since $|L(j\omega_c)| = 1$, the corresponding evaluation of the frequency response of the loop transfer function for $\omega = \omega_c$ results

$$L(j\omega_c) = |L(j\omega_c)|\, e^{-\angle L(j\omega_c)} = e^{-j\pi}$$

Hence, the closed-loop characteristic equation

$$1 + L(s) = 0$$

has a solution (i.e. a closed-loop pole) for $s = j\omega_c$. Since its real part is zero, the closed-loop system cannot be asymptotically stable.

6.3.2 Gain margin criterion

Similarly to the phase margin criterion, closed-loop asymptotic stability can be also verified with the following alternative way.

Theorem 8 (Gain margin criterion)**.** *Given that $L(s) = R(s)G(s)$ has no cancellations within the closed right half-plane and* ***no poles within the open right half-plane****, and assume*

1. *the generalised gain μ_L of $L(s)$ is positive, i.e. $\lim_{s\to 0} s^{g_L} L(s) > 0$;*
2. *the phase diagram of $L(j\omega)$ has* ***exactly one intersection*** *with the $-\pi$ axis from above to below occurring at frequency ω_π, i.e. $\exists!\omega_\pi : \angle L(j\omega_\pi) = -\pi$, $\angle L(j\omega) > -\pi, \forall\omega < \omega_\pi$ and $\angle L(j\omega) < 0, \forall\omega > \omega_\pi$.*

define g_m ***gain margin*** *the quantity*

$$g_m = \frac{1}{|L(j\omega_\pi)|}$$

then the closed-loop system of Fig. 6.8 is asymptotically stable if and only if

$$g_m > 1, \ \ i.e.\ \ |L(j\omega_\pi)| < 1$$

Repeating the same reasoning introduced for the phase margin, the value $g_m = 1$ corresponds to

$$|L(j\omega_\pi)| = 1$$

and since, from the definition of ω_π

$$\angle L(j\omega_\pi) = -\pi$$

we have

$$L(j\omega_\pi) = |L(j\omega_\pi)| e^{j\angle L(j\omega_\pi)} = e^{-j\pi}$$

from which the presence of a closed-loop pole in $s = j\omega_\pi$ follows, and consequently the impossibility for the closed-loop system to be asymptotically stable.

Example 36. *Considering again the following loop transfer function*

$$L(s) = \frac{5}{(1+s)^3}$$

which we already proved, using the Bode's criterion, to correspond to an asymptotically stable closed-loop system.

The Bode's diagrams of $L(j\omega)$ are reported in Fig. 6.11 from which we notice that all hypotheses for the applicability of the gain margin criterion are satisfied with $\omega_\pi \approx 1.73$ rad/s and $g_m \approx 1.6 > 1$, which further confirms the asymptotic stability of the corresponding closed-loop system.

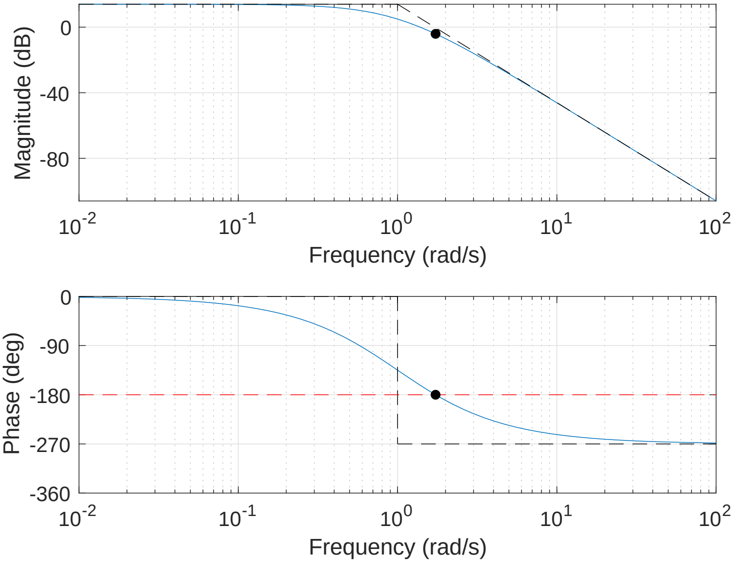


Figure 6.11: Bode's diagrams of $L(j\omega)$, from which $\omega_\pi \approx 1.73$ rad/s and $g_m \approx 1.6$.

Chapter 7

LTI SISO control systems analysis

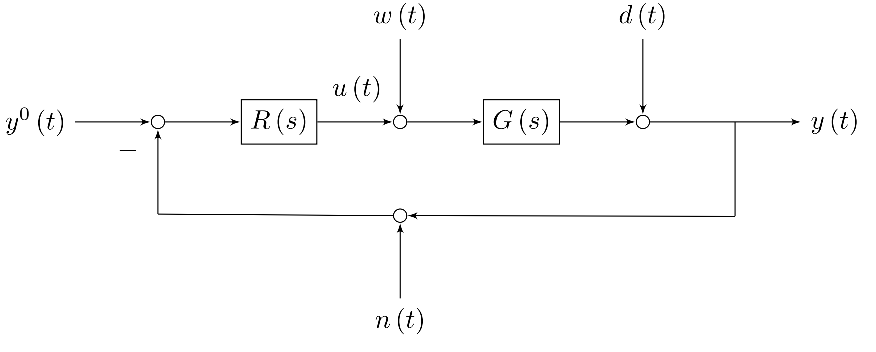


Figure 7.1: A negative feedback control system.

The focus of this Chapter is to provide tools for the analysis of the control system in Fig. 7.1, where

- $y^0(t)$ represents the exogenous signal which provides a (possibly time-varying) reference value to be tracked by the controlled variable $y(t)$;
- $w(t)$ represents a non-measurable load disturbance;
- $d(t)$ represents a non-measurable output disturbance;
- $n(t)$ represents the measurement noise;
- $u(t)$ represents the control variable.

In order to evaluate the tracking performance, we introduce the quantity (**error**):

$$e(t) = y^0(t) - y(t)$$

Finally, $R(s)$ and $G(s)$ are transfer functions representing the controller and the system to be controlled, respectively.

7.1 Stability of the closed-loop system

In Chapter 2, we have introduced stability as a property guaranteeing a bounded deviation from the nominal trajectory of a system in case of a perturbed initial condition. We have also introduced asymptotic stability in case of a vanishing effect of such a perturbation. We have also discussed that, in case of LTI systems, stability is a property that is shared among all possible trajectories. Still considering LTI systems, using the Lagrange's formula, we have seen asymptotic stability implies a vanishing free response of the state and of the corresponding output.

In Chapter 6 we have seen that the asymptotic stability for closed-loop system could be analysed through simple graphical arguments using the Bode's stability criterion, that is

- both the controller $R(s)$ and the system $G(s)$ **do not have** poles with strictly positive real part;
- the loop transfer function $L(s)$ has a positive gain;
- the magnitude diagram of the frequency response of the loop transfer function $L(s)$ crosses the 0 dB axis exactly ones from top to bottom;
- the phase margin is positive.

7.2 The "Gang of four"

The overall closed-loop system in Fig. 7.1 is characterised by four inputs ($y^0(t)$, $w(t)$, $d(t)$, and $n(t)$) and three outputs ($y(t)$, $u(t)$ and $e(t)$), see also Fig. 7.2. Assuming a null initial condition for all state variables, the relationships between each output variable and the three possible inputs are as follows:

$$\begin{aligned}
Y(s) &= \frac{L(s)}{1+L(s)}Y^0(s) + \frac{1}{1+L(s)}D(s) - \frac{L(s)}{1+L(s)}N(s) + \frac{G(s)}{1+L(s)}W(s) \\
E(s) &= \frac{1}{1+L(s)}Y^0(s) - \frac{1}{1+L(s)}D(s) - \frac{L(s)}{1+L(s)}N(s) - \frac{G(s)}{1+L(s)}W(s) \\
U(s) &= \frac{R(s)}{1+L(s)}Y^0(s) - \frac{R(s)}{1+L(s)}D(s) - \frac{R(s)}{1+L(s)}N(s) - \frac{L(s)}{1+L(s)}W(s)
\end{aligned} \tag{7.1}$$

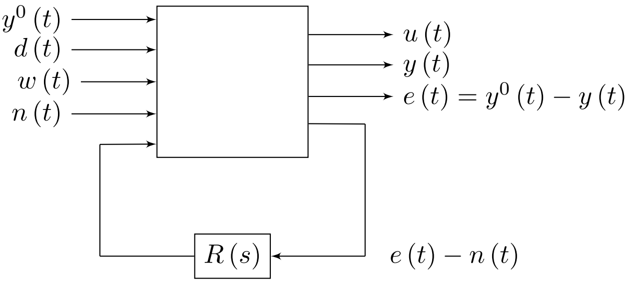


Figure 7.2: Pictorial representation of a negative feedback control system.

As one can notice from (7.1), despite the twelve possible pairs of inputs and outputs, the transfer functions for the closed-loop system are essentially four (apart from the sign), namely

1. the **complementary sensitivity**: $F(s) = \dfrac{L(s)}{1+L(s)}$;
2. the **sensitivity**: $S(s) = \dfrac{1}{1+L(s)}$;
3. the **control sensitivity**: $Q(s) = \dfrac{R(s)}{1+L(s)}$, and
4. the **output sensitivity**: $V(s) = \dfrac{G(s)}{1+L(s)}$.

which only differ from the numerator.

From now on, we assume that the loop transfer function $L(s) = R(s)G(s)$ characterising the closed-loop control system in Fig. 7.1 has been obtained without cancellations within the closed right half-plane, has a positive gain and no poles with strictly positive real part, it crosses the 0 dB axis exactly ones from top to bottom and $\psi_m > 0$. Under these assumptions, from the Bode's stability criterion, we can ensure asymptotic stability of the closed-loop system. As a consequence, the free motion of every state variable will tend to zero, and the frequency response theorem can be safely applied.

7.2.1 Complementary sensitivity

The complementary sensitivity expresses, among the others, the relationship between $y^0(t)$ and $y(t)$. Hence ideally one would like to have

$$\forall\omega : |F(j\omega)| = \left|\frac{L(j\omega)}{1+L(j\omega)}\right| \approx 1$$

which would imply $y(t) \approx y^0(t)$, hence $e(t) \approx 0$. To this end, one would require the magnitude of $L(j\omega)$ to be **as big as possible**.

On the other hand, notice that the same transfer function also describes the relationship between the measurement noise $n(t)$ and the controlled variable $y(t)$. For this reason, one would also like to obtain

$$\forall \omega : |F(j\omega)| = \left| \frac{L(j\omega)}{1 + L(j\omega)} \right| \approx 0$$

which would imply $y(t) \approx 0$, no matter the amplitude of the measurement noise $n(t)$. To this end, one would require the magnitude of $L(j\omega)$ to be **as small as possible**.

The two requirements are clearly **conflicting**. Moreover, since the loop transfer function $L(s)$ is assumed to satisfy the Bode's stability criterion, its magnitude diagram crosses the 0 dB axis exactly once, from top to bottom, corresponding to the crossover frequency $\omega = \omega_c$. Hence, a reasonable requirement is to have $|L(j\omega)|$ as big as possible for $\omega \ll \omega_c$ as well as $|L(j\omega)|$ as small as possible for $\omega \gg \omega_c$. Consistently, the complementary sensitivity would result

$$|F(j\omega)| \approx \begin{cases} 1, \omega \ll \omega_c \\ |L(j\omega)|, \omega \gg \omega_c \end{cases} \tag{7.2}$$

Example 37. *[Approximations for the complementary sensitivity] In one examples of Chapter 6, we considered*

$$G(s) = \frac{1}{1+s}, R(s) = 1000\frac{1+0.1s}{s}$$

and the corresponding loop transfer function

$$L(s) = \frac{1000}{s}\frac{1+0.1s}{1+s}$$

Correspondingly, we have computed a crossover frequency of $\omega_c \approx 100$ rad/s and a phase margin of $\psi_m \approx 85$ deg. Figure 7.3 reports the magnitude diagrams of $L(j\omega)$, $F(j\omega)$. As one can see, the approximation introduced in (7.2) well predicts the behaviour of the actual magnitude of the frequency response of $F(s)$.

As of now, we have been able to approximate the magnitude diagram of $F(j\omega)$ at frequencies that are well separated from the crossover frequency ω_c, namely for $\omega \ll \omega_c$ and $\omega \gg \omega_c$. What happens in the neighbourhood of ω_c is still to be investigated. In particular, since $|L(j\omega_c)| = 1$, we have

$$|F(j\omega_c)| = \frac{|L(j\omega_c)|}{|1 + L(j\omega_c)|} = \frac{1}{|1 + L(j\omega_c)|}$$

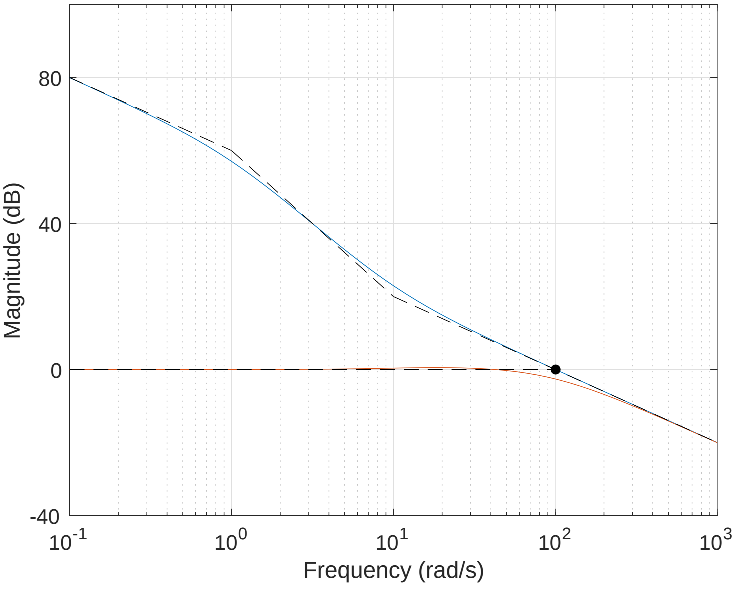


Figure 7.3: Bode's diagrams of the magnitude of $L\left(j\omega\right)$ (blue), $F\left(j\omega\right)$ (red).

Moreover, as $L\left(j\omega_c\right) = e^{j\angle L\left(j\omega_c\right)} = e^{j\psi_c} = \cos\left(\psi_c\right) + j\sin\left(\psi_c\right)$, one obtains

$$\left|F\left(j\omega_c\right)\right| = \frac{1}{\sqrt{\left(1+\cos\left(\psi_c\right)\right)^2 + \sin\left(\psi_c\right)^2}} = \frac{1}{\sqrt{2\left(1+\cos\left(\psi_c\right)\right)}}$$

We can relate the phase of the loop transfer function with the corresponding phase margin as $\psi_c = \psi_m - \pi$ and therefore $\cos\left(\psi_c\right) = -\cos\left(\psi_m\right)$. Finally, from half-angle formulas and since we assumed $\psi_m > 0$

$$\left|F\left(j\omega_c\right)\right| = \frac{1}{2\sin\left(\dfrac{\psi_m}{2}\right)} \tag{7.3}$$

It is evident that the behaviour of $\left|F\left(j\omega_c\right)\right|$ is highly sensitive to the value of the phase margin.

Example 38. *[Behaviour of the complementary sensitivity around ω_c] Consider the two following loop transfer functions*

$$L_1\left(s\right) = \frac{0.5}{s}, L_2\left(s\right) = \frac{1}{s\sqrt{2}\left(1+s/0.5\right)}$$

which are characterised by exactly the same crossover frequency $\omega_c = 0.5$ rad/s, but different phase margins, $\psi_m = 90$ deg and $\psi_m = 45$ deg, respectively.

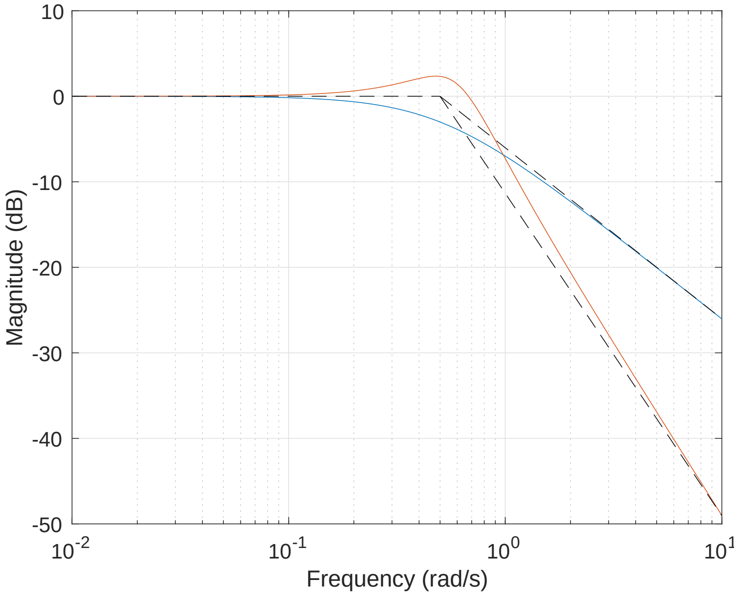


Figure 7.4: Magnitude diagrams of the frequency response of the two complementary sensitivities $F_1(j\omega)$ (blue), $F_2(j\omega)$ (red).

The magnitude diagrams of the frequency responses of the corresponding complementary sensitivities are reported in Fig. 7.4.

While the former, corresponding to $\psi_m = 90$ deg*, has a magnitude lower than* 0 *dB corresponding to the crossover frequency* $\omega_c = 0.5$ *rad/s, the magnitude of the latter is definitely higher than* 1*. This could have been somehow expected from* (7.3)*. In particular, the shape of* $|F_1(j\omega)|$ *resembles the one of a first-order system with unit gain and pole in* $1/\omega_c$*, while the one of* $|F_2(j\omega)|$ *presents a resonance peak which can only be justified by a pair of complex and conjugate poles.*

Based on the evidence of the previous example, depending on the value of the phase margin ψ_m, the behaviour of the complementary sensitivity can be further approximated. A first-order approximation

$$F(s) \approx \frac{1}{1 + s/\omega_c} \tag{7.4}$$

can be justified in case of high values of the phase margin (for example, $\psi_m > 80$ deg), while, in the case of significantly lower values (i.e. $\psi_m < 40$ deg), a second-order approximation

$$F(s) \approx \frac{1}{s^2/\omega_c^2 + 2\xi/\omega_c s + 1}$$

should be preferred. Further notice that the position of the pole(s) coincides with the crossover frequency ω_c. Moreover, the value of the damping ξ in the second-order approximation deserves further attention. From the analysis of Bode's diagrams in Chapter 3, we have seen that the peak magnitude of a resonant behaviour is equal to $1/(2\xi)$. On the other hand, we have already seen in (7.3) that the value $|F(j\omega_c)|$ depends on the phase margin ψ_m, hence letting

$$\xi = \sin\left(\frac{\psi_m}{2}\right) \tag{7.5}$$

We can conclude that, in case of reduced phase margins, an approximated version of the complementary sensitivity can be the following one:

$$F(s) \approx \frac{1}{\frac{s^2}{\omega_c^2} + 2\frac{\sin(\psi_m/2)}{\omega_c}s + 1} \tag{7.6}$$

Finally, assuming the measurement noise $n(t)$ to be the combination of high frequency harmonic signals with frequencies $\bar{\omega} \gg \omega_c$, from the frequency response theorem we have that

$$\forall \bar{\omega} \gg \omega_c : |Y(j\bar{\omega})| = |F(j\bar{\omega})|\,|N(j\bar{\omega})| \approx |L(j\bar{\omega})|\,|N(j\bar{\omega})|$$

meaning that effect of the measurements noise $n(t)$ on the controlled output $y(t)$ will be attenuated since $|L(j\bar{\omega})| \ll 1$ for $\bar{\omega} \gg \omega_c$.

Example 39. *[Effect of measurement noise on the output] Consider*

$$G(s) = \frac{1}{1+s}$$

and two different controllers

$$R_1(s) = 1000\frac{1+0.1s}{s}, R_2(s) = 1000\frac{1+0.1s}{s(1+0.001s)}$$

where the latter is identical to the former, apart from an additional high-frequency pole. In both the cases, $\omega_c = 100$ rad/s with $\psi_m > 0$.
Figure 7.5 reports the magnitude diagrams of the frequency responses of the two complementary sensitivites $F_1(s)$ and $F_2(s)$ corresponding to the two different controllers. Assume now, that the closed-loop system is affected by a high frequency measurement noise $n(t) = \sin(\bar{\omega}t)$ with $\bar{\omega} = 5000$ rad/s. Since the magnitudes of the two corresponding frequency responses of the complementary sensitivity differ of approximately 15 dB at the frequency of the measurement noise $\bar{\omega} = 5000$ rad/s, we expect the amplitude of the oscillations in the output signal $y(t)$ corresponding to the first controller to be $10^{(15/20)} \approx 5.6$ higher than in the other case (see Fig.7.6).

7.2.2 Sensitivity

Back to (7.1), the sensitivity transfer function describes the relationship between the disturbance $d(t)$ and the error $e(t)$, as well as between the disturbance $d(t)$ or the reference $y^0(t)$ and the error $e(t)$. In all the cases, one would

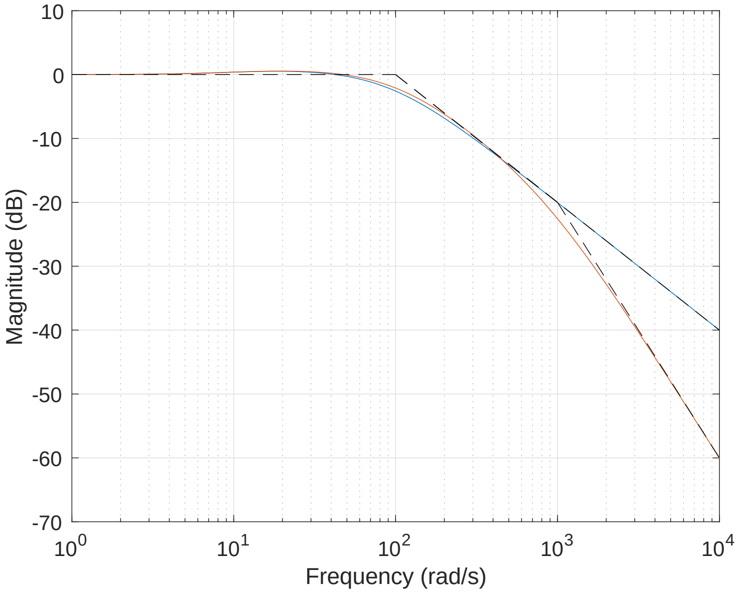


Figure 7.5: Magnitude diagrams of the frequency responses of the two complementary sensitivities $F_1(j\omega)$ (blue), $F_2(j\omega)$ (red).

like to have its magnitude as small as possible, at all frequencies, i.e.

$$\forall\omega : |S(j\omega)| = \left|\frac{1}{1+L(j\omega)}\right| \approx 0$$

This would imply the output $y(t)$, as well as the tracking error $e(t)$, being insensitive to the disturbance $d(t)$. Unfortunately, the following relationship holds

$$\forall s : F(s) + S(s) = 1$$

and therefore the magnitude of the frequency response of the sensitivity transfer function can be surely made small at frequencies lower than ω_c, but it will inevitably get close to 0 dB for frequencies higher than the crossover frequency, since $|F(j\omega)| \approx 0, \omega \gg \omega_c$. As we have already seen discussing the complementary sensitivity, a reasonable requirement is to have $|L(j\omega)|$ as big as possible for $\omega \ll \omega_c$ and as small as possible for $\omega \gg \omega_c$. Consistently, the sensitivity would result

$$|S(j\omega)| \approx \begin{cases} 1/|L(j\omega)|, \omega \ll \omega_c \\ 1, \omega \gg \omega_c \end{cases} \tag{7.7}$$

Example 40. *[Approximations of the sensitivity] Considering again*

$$L(s) = \frac{1000}{s}\frac{1+0.1s}{1+s}$$

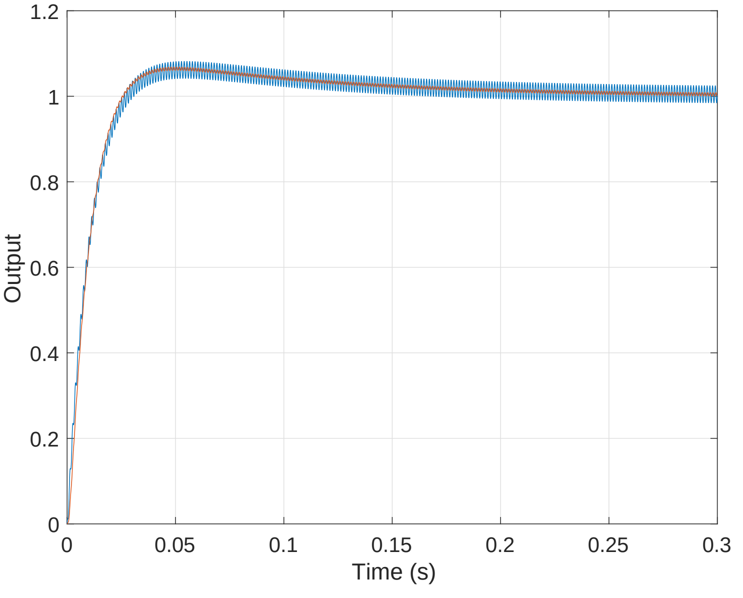


Figure 7.6: Responses of the output signal $y(t)$ for the two controllers $R_1(s)$ (blue) and $R_2(s)$ (red).

the corresponding magnitudes of the frequency responses of the sensitivity and the complementary sensitivity transfer functions are reported in Fig. 7.7. As expected, since $F(s) + S(s) = 1$, wherever $|F(j\omega)| \approx 1$ (i.e. for $\omega \ll \omega_c$), $|S(j\omega)| = 1/|L(j\omega)| \approx 0$, while corresponding to $|F(j\omega)| = |L(j\omega)| \approx 0$ (i.e. when $\omega \gg \omega_c$), $|S(j\omega)| \approx 1$.

We now wonder why transfer function $S(s)$ is called sensitivity. To understand the reason behind its name, consider the complementary sensitivity

$$F(s) = \frac{R(s)\,G(s)}{1 + R(s)\,G(s)}$$

and assume one may want to investigate how this closed-loop transfer function is affected by a variation of the transfer function $G(s)$ representing the system under control. In other words, one wants to study $\partial F(s)/\partial G(s)$. With simple mathematical arguments, we obtain

$$\frac{\partial F(s)}{\partial G(s)} = \frac{\partial}{\partial G(s)}\left(\frac{R(s)\,G(s)}{1 + R(s)\,G(s)}\right) = \frac{R(s)}{(1 + R(s)\,G(s))^2} = \frac{1}{1 + R(s)\,G(s)}\frac{F(s)}{G(s)}$$

from which

$$\frac{\partial F(s)}{F(s)} = S(s)\,\frac{\partial G(s)}{G(s)} \tag{7.8}$$

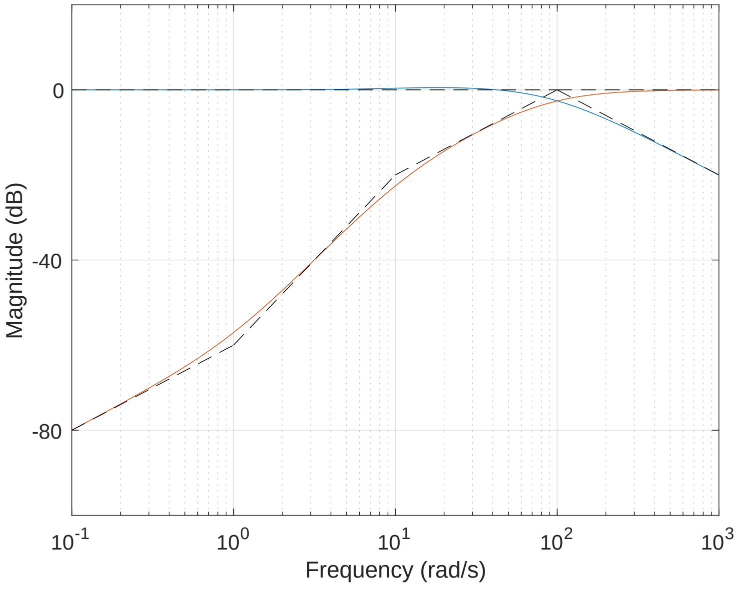


Figure 7.7: Magnitude diagrams of the frequency responses of the complementary sensitivity $F(j\omega_c)$ (blue) and of the sensitivity $S(j\omega_c)$ (red) transfer functions.

From equation (7.8), one can notice that a (normalised) variation in the frequency response of the system under control, i.e. $\partial G(j\omega)/G(j\omega)$, is producing a (normalised) variation in the frequency response complementary sensitivity, $\partial F(j\omega)/F(j\omega)$, only when $S(j\omega)$ differs from zero, which happens for $\omega \gg \omega_c$. We can then conclude that the feedback control scheme in Fig. 7.1 is able to mask the effects of uncertainties $\partial G(j\omega)$, provided that they happen before the crossover frequency ω_c, i.e. when $|S(j\omega)| \approx 0$.

Example 41. *[Mass with friction] Consider again the system in Fig. 6.1. Assume the velocity of the mass $v(t)$ as the measured output of the system, $m = 50$, $r = 3$, the transfer function from the input force $f(t)$ to the velocity $v(t)$ is as follows*

$$G(s) = \frac{1}{50s + 3}$$

Further assume a controller $R(s) = K$ with $K = 500$, then the corresponding loop transfer function

$$L(s) = \frac{500}{50s + 3}$$

has a crossover frequency of $\omega_c = 10$ rad/s and a phase margin of $\psi_m \approx 90$ deg.

Assume that the friction coefficient of the actual system is ten times smaller the the one assumed so far. In this case

$$G(s) = \frac{1}{50s + 0.3}$$

and

$$L(s) = \frac{500}{50s + 0.3}$$

which still corresponds to $\omega_c = 10$ *rad/s and* $\psi_m \approx 90$ deg. *Figure 7.8 reports the magnitudes of the frequency responses of the two loop transfer functions.*

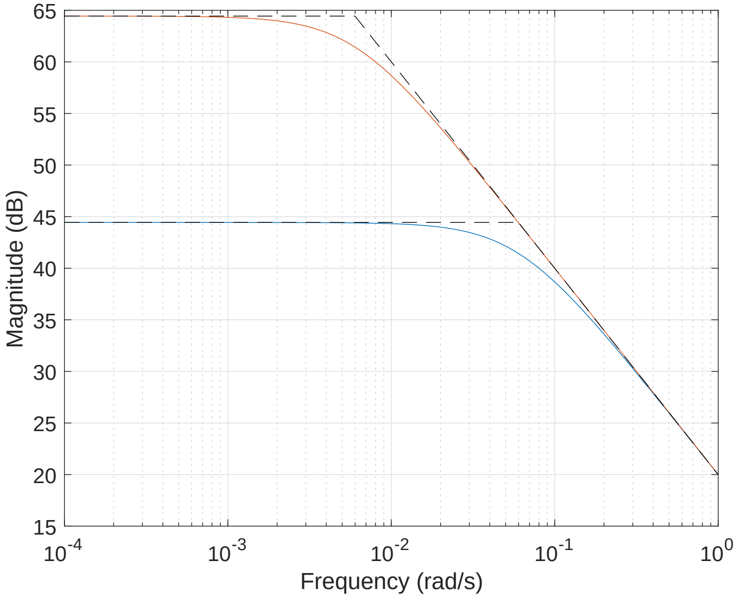


Figure 7.8: Magnitude diagrams of the frequency responses of the two loop transfer functions for the nominal (blue) and the reduced (red) friction coefficient.

As one can see, the major difference between the two (which is due to the two different values of the friction coefficients r*) happens way before the crossover frequency* ω_c.

One would then expect the two corresponding complementary sensitivities to be roughly the same. Indeed, Fig. 7.9 reports the magnitude diagrams of the frequency responses of the two complementary sensitivities. As the uncertainty on the dynamics of the system (and therefore on the corresponding loop transfer function) happens at low frequency, i.e. for $\omega \ll \omega_c$*, the modification in the friction coefficient has barely no effects on the behaviour of the closed-loop system and in particular on the response of* $v(t)$ *to a constant reference velocity* $v^0(t) = \bar{v}, t \geq 0$. *In Chapter 6, though with different arguments, we were able to get to the same insights, in fact for* $R(s) = K$*, we concluded that the forced response of the output* $v(t)$ *was*

$$v(t) = \frac{K\bar{v}}{K+r}\left(1 - e^{-\frac{K+r}{m}t}\right), t \geq 0$$

which is barely influenced by r*, in case* $K \gg r$.

In Chapter 6, as well as in the last example, we have seen that the controller $R(s) = K$ was able to make the velocity of the mass $v(t)$ arbitrarily close to

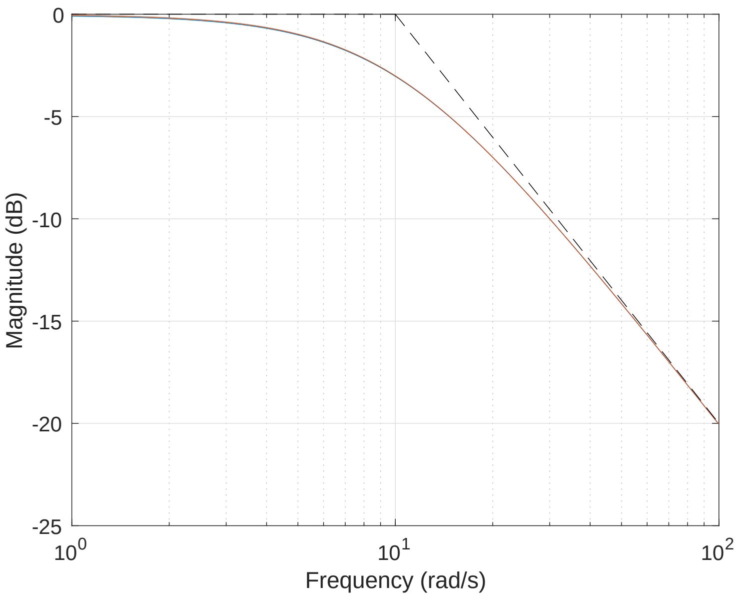


Figure 7.9: Magnitude diagrams of the frequency responses of the two complementary sensitivities for the nominal (blue) and the reduces value of the friction coefficient (red).

the desired value $\bar{v}$, but it was not possible to make it exactly equal to the reference, at least asymptotically.

More generally, we are interested in formally investigating which property of the control system is influencing the steady-state value of the error $e(t) = y^0(t) - y(t)$. As already discussed, the sensitivity transfer function $S(s)$ describes how the error signal $e(t)$ is related to both the reference $y^0(t)$ and the disturbance $d(t)$. And therefore, assuming $w(t) = n(t) = 0$, we have

$$e_\infty = \lim_{t \to +\infty} e(t) = \lim_{s \to 0} sE(s) = \lim_{s \to 0} sS(s)\left(Y^0(s) - D(s)\right)$$

Without any lack of generality we can assume $d(t) = 0$ and analyse the effects of $y^0(t)$, only. The following reasoning can be then applied considering $d(t)$ and $y^0(t) = 0$. Considering a reference signal $y^0(t)$ as a polynomial function of time t, that is

$$y^0(t) = t^k, k \geq 0, t \geq 0$$

we have

$$e_\infty = \lim_{s \to 0} sS(s)s^{-k-1} = \lim_{s \to 0} S(s)s^{-k}$$

which clearly depends on k and on the behaviour of $S(s)$ for $s \to 0$. The latter clearly depends on the behaviour of $L(s)$ for $s \to 0$ which satisfies

$$\lim_{s \to 0} L(s) - \frac{\mu_L}{s^{g_L}} = 0$$

hence

$$e_\infty = \lim_{s\to 0} \frac{s^{-k}}{1+\dfrac{\mu_L}{s^{g_L}}} = \lim_{s\to 0} \frac{s^{g_L-k}}{s^{g_L}+\mu_L} = \begin{cases} \dfrac{1}{1+\mu_L}, g_L = k = 0 \\ \dfrac{1}{\mu_L}, g_L = k, g_L > 0 > 0 \\ 0, g_L > k \\ +\infty, g_L < k \end{cases}$$

From the time domain perspective, the Tab. 7.1 summarises the given findings:

g_L	$y^0(t) = 1, t \geq 0$	$y^0(t) = t, t \geq 0$	$y^0(t) = t^2, t \geq 0$	$\dots$
0	$\dfrac{1}{1+\mu_L}$	$+\infty$	$+\infty$	$\dots$
1	0	$\dfrac{1}{\mu_L}$	$+\infty$	$\dots$
2	$\vdots$	0	$\dfrac{1}{\mu_L}$	$\ddots$
$\vdots$	$\vdots$	$\vdots$	0	$\ddots$

Table 7.1: Steady-state error e_∞ as function of the type g_L and generalised gain μ_L of the loop transfer, for some possible reference signals $y^0(t)$.

Hence, the type g_L and the gain μ_L of the loop transfer function $L(s)$ are important parameters influencing the precision of the control system in the asymptotic tracking of the reference.

Example 42. *[Mass with friction] With reference again to the simple system of Fig. 6.1, in Chapter 6 we have seen that for a constant reference velocity $v^0(t) = \bar{v}, t \geq 0$, no disturbance $w(t) = 0, \forall t \geq 0$, and controller $R(s) = K$, the steady-steady state value of the velocity could be characterised as follows*

$$v_\infty = \bar{v}\frac{K}{K+r}$$

In the light of the discussion above on the sensitivity transfer function, we would have obtained the same results simply computing

$$v_\infty = \bar{v} - e_\infty = \bar{v}\left(1 - \lim_{s\to 0} S(s)\right) = \bar{v}\left(1 - \frac{1}{1+K/r}\right) = \bar{v}\frac{K}{K+r}$$

On the other hand, with the controller

$$R(s) = 500\frac{1+s}{s}$$

hence with a loop transfer function with type $g_L = 1$, we would have obtained

$$v_\infty = \bar{v} - e_\infty = \bar{v}\left(1 - \lim_{s\to 0} S(s)\right) = \bar{v}\left(1 - \lim_{s\to 0}\frac{1}{1+\dfrac{500}{rs}}\right) = \bar{v}$$

Consistently with the findings in Tab. 7.1, a different type g_L of the loop transfer function is responsible for a different steady-state behaviour of the output in tracking a constant reference.

The sensitivity transfer function also represents the relationship between the disturbance $d(t)$ and the output $y(t)$. Ideally one would like to have the output insensitive to the disturbance, we have seen that this is only possible at frequencies lower than the crossover frequency, i.e. for $\omega \ll \omega_c$. Then, assuming the disturbance $d(t)$ to contain just frequencies $\bar{\omega}$ that are lower than ω_c, i.e. $\bar{\omega} \ll \omega_c$, from the frequency response theorem we have that

$$\forall \bar{\omega} \ll \omega_c : |E(j\bar{\omega})| = |S(j\bar{\omega})| \, |D(j\bar{\omega})| \approx |1/L(j\bar{\omega})| \, |D(j\bar{\omega})|$$

meaning that the effect of the disturbance $d(t)$ on the output $y(t)$ will be attenuated by a factor that depends on how large $|L(j\bar{\omega})|$ is.

Example 43. *[Effect of the disturbance on the output] Consider a system and two different controller*

$$G(s) = \frac{1}{1+s}, R_1(s) = 1000\frac{1+0.1s}{s}, R_2(s) = 100\frac{1+s}{s}$$

For both the controllers, $\omega_c = 100$ rad/s with $\psi_m > 0$.

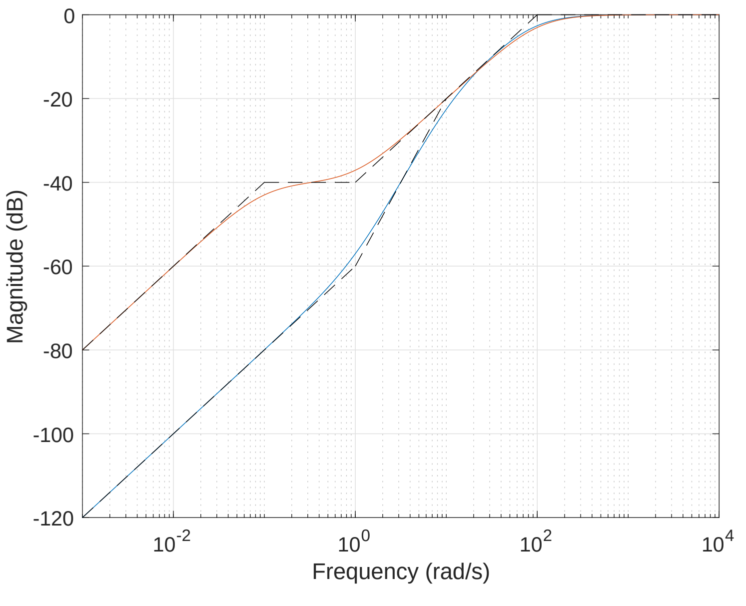


Figure 7.10: Magnitude diagrams of the frequency responses of the two sensitivities $S_1(s)$ (blue) and $S_2(s)$ (red).

Figure 7.10 reports the magnitude diagrams of the frequency responses of the two sensitivities $S_1(s)$ and $S_2(s)$ corresponding to the two different controllers $R_1(s)$ and $R_2(s)$.

Assume now, that the closed-loop system is affected by a slowly-varying disturbance $d(t) = \sin(\bar{\omega} t)$ with $\bar{\omega} = 1\ rad/s \ll \omega_c = 100\ rad/s$. Since the magnitudes of the two corresponding frequency responses of the sensitivity differ of approximately 20 dB at the frequency of the disturbance $\bar{\omega} = 1\ rad/s$, we expect the amplitude of the oscillations in the output signal $y(t)$ corresponding to the first controller to be 10 times smaller than in the other case (see Fig.7.11).

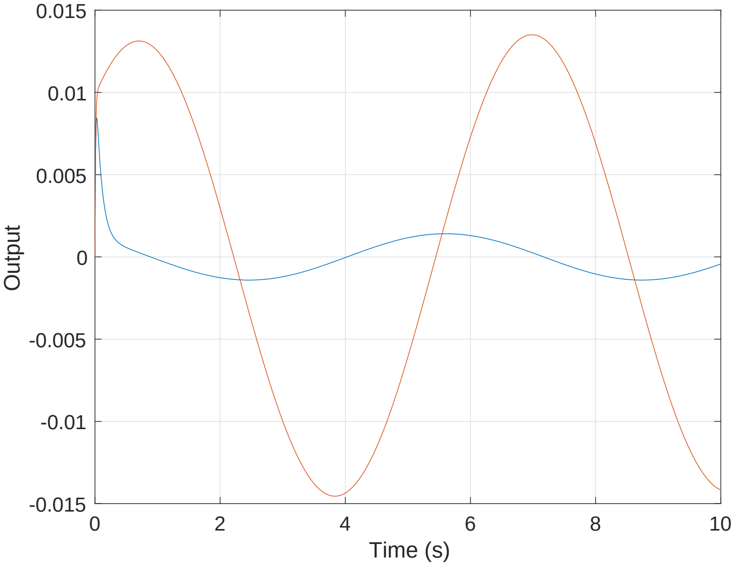


Figure 7.11: Responses of the output signal $y(t)$ for the two controllers $R_1(s)$ (blue) and $R_2(s)$ (red).

7.2.3 Control sensitivity

From (7.1), we see that the control sensitivity

$$Q(s) = \frac{R(s)}{1 + L(s)}$$

equivalently describes, apart from the sign, how the control variable $u(t)$ is related to the reference $y^0(t)$, the disturbance $d(t)$, and the measurement noise $n(t)$. Therefore, one would ideally like to have

$$\forall \omega : |Q(j\omega)| = \left| \frac{R(j\omega)}{1 + L(j\omega)} \right| \approx 0$$

in order to achieve all possible closed-loop performance, with negligible control effort. We should now have understood that ideal requirements cannot be satisfied in practice. In a more realistic situation of $|L(j\omega)|$ being big for $\omega \ll \omega_c$ and small for $\omega \gg \omega_c$, one might introduce the following approximations:

$$|Q(j\omega)| \approx \begin{cases} 1/|G(j\omega)|, \omega \ll \omega_c \\ |R(j\omega)|, \omega \gg \omega_c \end{cases} \tag{7.9}$$

For the first time, the approximation of one transfer function in the "Gang of four" does not depend directly on the magnitude of $L(j\omega)$. From the one hand, at frequencies lower than the crossover frequency, i.e. $\omega \ll \omega_c$, the magnitude of the frequency response of the control sensitivity depends solely on the system to be controlled.

On the other hand, at higher frequencies, that is for $\omega \gg \omega_c$, the magnitude of the frequency response of the controller $|R(j\omega)|$ might significantly influence how the control signal $u(t)$ is affected by the reference $y^0(t)$, the disturbance $d(t)$, and the measurement noise $n(t)$.

Example 44. *[Approximations of the control sensitivity] Consider*

$$G(s) = \frac{1}{1+s}, R(s) = 1000\frac{1+0.1s}{s}$$

The loop transfer function corresponds to a crossover frequency of $\omega_c = 100$ rad/s and a positive phase margin, i.e. $\psi_m > 0$.

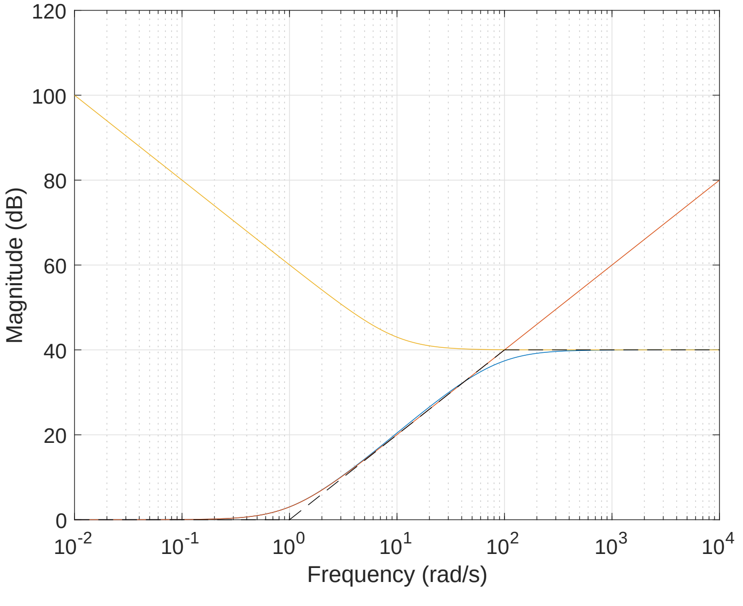


Figure 7.12: Magnitude diagram of the frequency response of the control sensitivity $Q(s)$ (blue) as compared to the ones of $1/G(s)$ (red) and $R(s)$ (yellow).

Figure 7.12 reports the Bode's diagram of the magnitude of $Q(j\omega)$ as compared to the magitudes of $G(j\omega)$ and $R(j\omega)$.

Finally, assuming the measurement noise $n(t)$ to be the combination of high frequency harmonic signals, i.e. with frequencies $\bar{\omega} : \bar{\omega} \gg \omega_c$, from the frequency response theorem we have that

$$\forall \bar{\omega} \gg \omega_c : |U(j\bar{\omega})| = |Q(j\bar{\omega})| \, |N(j\bar{\omega})| \approx |R(j\bar{\omega})| \, |N(j\bar{\omega})|$$

meaning that the amplitudes of the harmonics of the measurements noise $n(t)$ will be possibly amplified in the control signal $u(t)$ for $\bar{\omega} \gg \omega_c$.

Example 45. *[Effect of measurement noise on control signal] Consider again*

$$G(s) = \frac{1}{1+s}$$

and two different controllers

$$R_1(s) = 1000\frac{1+0.1s}{s}, R_2(s) = 1000\frac{1+0.1s}{s(1+0.001s)}$$

where the latter is identical to the former, apart from an additional high-frequency pole. Figure 7.13 reports the magnitude diagrams of the frequency responses of the two control sensitivites $Q_1(s)$ and $Q_2(s)$ corresponding to the two different controllers $R_1(s)$ and $R_2(s)$.

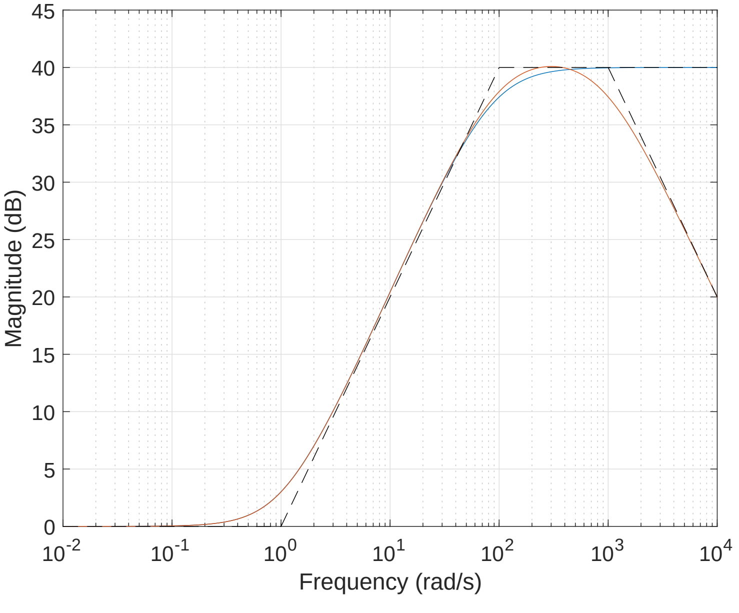


Figure 7.13: Magnitude diagrams of the frequency responses of the two control sensitivities corresponding to controller $R_1(s)$ (blue) and $R_2(s)$ (red).

Assume now that the closed-loop system is affected by a high frequency measurement noise $n(t) = 0.1\sin(\bar{\omega}t)$ with $\bar{\omega} = 5000$ rad/s $\gg \omega_c = 1000$ rad/s. Since the magnitudes of *the two corresponding frequency responses of the control sensitivity differ of approximately 15 dB at the frequency of the measurement noise, i.e. for $\bar{\omega} = 5000$ rad/s, we expect the amplitude of the oscillations in the control signal $u(t)$ corresponding to the first controller to be $10^{(15/20)} \approx 5.6$ times higher than in the other case (see Fig. 7.14).*

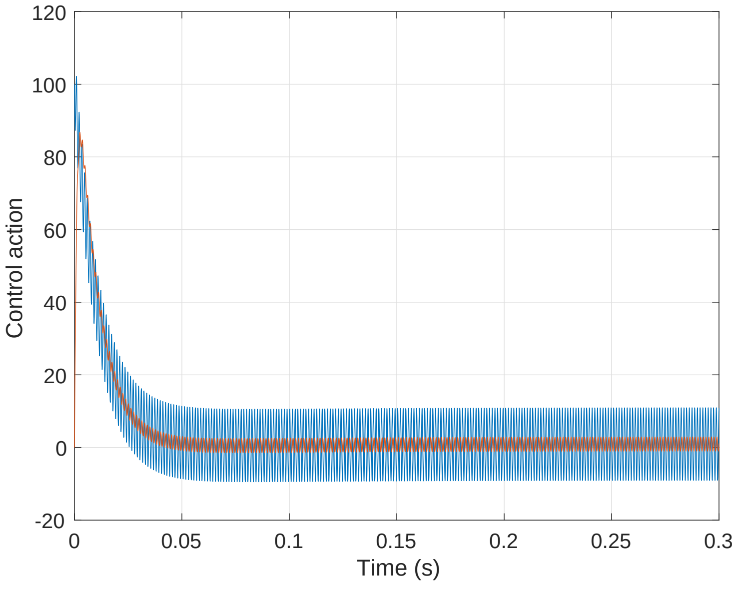


Figure 7.14: Responses of the control signal $u(t)$ to the measurement noise $n(t)$ for the two controllers $R_1(s)$ (blue) and $R_2(s)$ (red), respectively.

7.2.4 Output sensitivity

Last but not least, the output sensitivity $V(s)$ expresses the relationship between the load disturbance $w(t)$ and the error $e(t)$. Ideally, one would like to have

$$\forall\omega : |V(j\omega)| = \left|\frac{G(j\omega)}{1 + L(j\omega)}\right| \approx 0$$

in order to make the error $e(t)$ insensitive to the load disturbance $w(t)$. However, as we have already seen, a more realistic requirement is to have $|L(j\omega)|$ big only before the crossover frequency, i.e. for $\omega \ll \omega_c$, in this case

$$|V(j\omega)| \approx \begin{cases} |G(j\omega)| / |L(j\omega)|, \omega \ll \omega_c \\ |G(j\omega)|, \omega \gg \omega_c \end{cases} \tag{7.10}$$

Therefore, assuming the load disturbance $w(t)$ to be characterised by low-frequency harmonics only, i.e. with frequencies $\bar{\omega} \ll \omega_c$, from the frequency response theorem we have that

$$\forall \bar{\omega} \ll \omega_c : |E(j\bar{\omega})| = |V(j\bar{\omega})| \, |W(j\bar{\omega})| \approx \frac{|G(j\bar{\omega})|}{|L(j\bar{\omega})|} |W(j\bar{\omega})|$$

Example 46. *[Effects of load disturbance on error] Consider one time more*

$$G(s) = \frac{1}{1+s}$$

and two different controllers

$$R_1(s) = 1000\frac{1+0.1s}{s}, R_2(s) = 10\frac{1+10s}{s}$$

Figure 7.15 reports the magnitude diagrams of the frequency responses of the two output sensitivities $V_1(s)$ and $V_2(s)$ corresponding to the two different controllers $R_1(s)$ and $R_2(s)$, respectively.

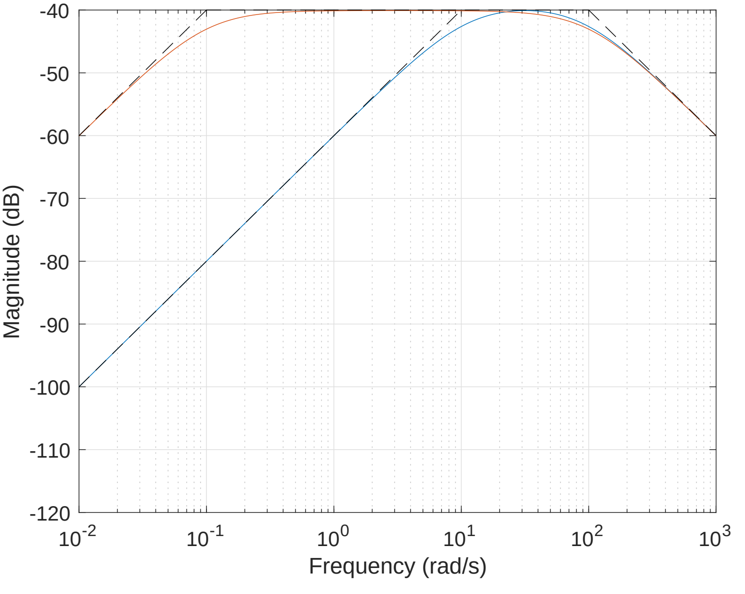


Figure 7.15: Magnitude diagrams of the frequency responses of the two output sensitivities corresponding to controllers $R_1(s)$ (blue) and $R_2(s)$ (red).

Assume now, that the closed-loop system is affected by a lo frequency load disturbance $w(t) = \sin(\bar{\omega}t)$ with $\bar{\omega} = 1$ rad/s $\ll \omega_c = 100$ rad/s. Since the magnitudes of the two corresponding frequency responses of the control sensitivities differ approximately of 20 dB at the frequency of the load disturbance $\bar{\omega} = 1$ rad/s, we expect the amplitude of the oscillations in the error signal $e(t)$ corresponding to the first controller to be 10 times higher than in the other case (see Fig.7.16).

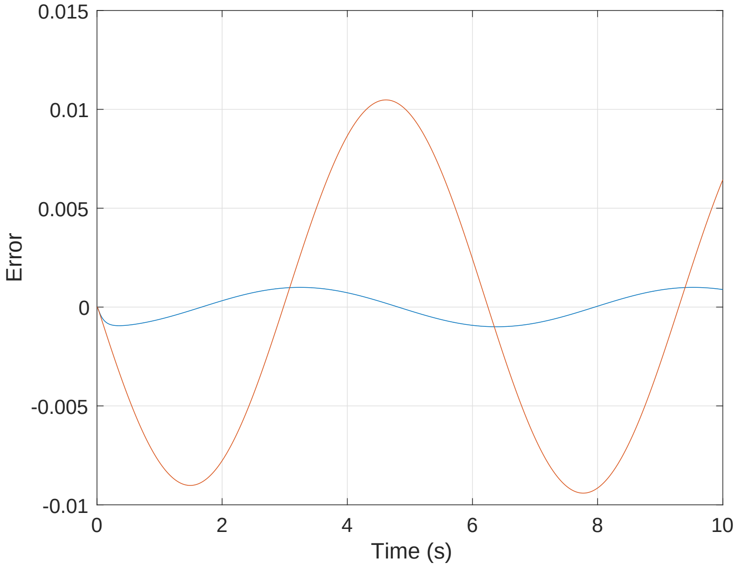


Figure 7.16: Responses of the error signal $e\left(t\right)$ to the load disturbance $w\left(t\right)$ for the two controllers $R_1\left(s\right)$ (blue) and $R_2\left(s\right)$ (red).

Finally, we might be interested in investigating which property of the control system influences the steady-state value of the error $e\left(t\right) = y^0\left(t\right) - y\left(t\right)$. Assuming $y^0\left(t\right) = d\left(t\right) = n\left(t\right) = 0$, we have

$$e_\infty = \lim_{t\to+\infty} e\left(t\right) = \lim_{s\to 0} sE\left(s\right) = -\lim_{s\to 0} sV\left(s\right)W\left(s\right)$$

Considering

$$w\left(t\right) = t^k, k \geq 0, t \geq 0$$

we have

$$e_\infty = -\lim_{s\to 0} sV\left(s\right)s^{-k-1} = -\lim_{s\to 0} V\left(s\right)s^{-k}$$

which clearly depends on k and on the behaviour of $V\left(s\right)$ for $s \to 0$. The latter clearly depends on the behaviour of both $L\left(s\right)$ and $G\left(s\right)$ for $s \to 0$, which satisfy

$$\lim_{s\to 0} L\left(s\right) - \frac{\mu_L}{s^{g_L}} = 0, \lim_{s\to 0} G\left(s\right) - \frac{\mu_G}{s^{g_G}} = 0$$

hence

$$\left|e_\infty\right| = \lim_{s\to 0} \frac{\frac{\mu_G}{s^{g_G}}}{1 + \frac{\mu_L}{s^{g_L}}} s^{-k} = \lim_{s\to 0} \frac{\mu_G s^{g_R - k}}{s^{g_L} + \mu_L}$$

From the time domain perspective, the Tab. 7.2 summarises the given findings:

g_R	$w(t)=1, t\geq 0$	$w(t)=t, t\geq 0$	$w(t)=t^2, t\geq 0$	$\ldots$
0	$\lim_{s\to 0} \dfrac{\mu_G}{s^{g_G}+\mu_L}$	$+\infty$	$+\infty$	$\ldots$
1	0	$\dfrac{1}{\mu_R}$	$+\infty$	$\ldots$
2	$\vdots$	0	$\dfrac{1}{\mu_R}$	$+\infty$
$\vdots$	$\vdots$	$\vdots$	0	$\dfrac{1}{\mu_R}$

Table 7.2: Steady-state error $|e_\infty|$ as function of the types g_L, g_G and generalised gain μ_L, μ_G of the loop transfer and of the system, for some possible reference disturbances $w(t)$.

Apart from the case $g_R = 0$ and $k = 0$ for which the steady-state error e_∞ might depend on the type g_G of the system, and in particular

$$e_\infty = \begin{cases} \dfrac{\mu_G}{1+\mu_R} & g_G = 0 \\ \dfrac{1}{\mu_R} & g_G > 0 \end{cases}$$

all the the other cases are very similar to those in Tab. 7.1 provided that μ_L is replaced with $\mu_R = \mu_L/\mu_G$.

Example 47. *[Mass with friction] Consider once again the simple system of Fig. 6.1, in Chapter 6 we have seen that for a constant load disturbance $w(t) = W, \forall t \geq 0$, null reference velocity $v^0(t) = 0$, and controller*

$$R(s) = K$$

the steady-steady state value of the velocity could be characterised as

$$e_\infty = -W\frac{1}{K+r}$$

In the light of the discussion above on the output sensitivity transfer function, we would have obtained the same results simply computing

$$e_\infty = \lim_{s\to 0} sE(s) = -W \lim_{s\to 0} V(s) = -\frac{W}{K+r}$$

On the other hand, with the controller

$$R(s) = 500\frac{1+s}{s}$$

hence with a controller and loop transfer function with type $g_R = g_L = 1$, we would have obtained

$$e_\infty = \lim_{s\to 0} sE(s) = -W \lim_{s\to 0} V(s) = 0$$

7.3 Summary

In this Chapter, we have analysed each of the transfer function in the "Gang of four". We are about to sum up all the findings but now relating them to each possible output of interest in the closed-loop system of Fig. 7.1, that is the output signal $y(t)$, the error signal $e(t)$, and the control signal $u(t)$.

The output signal $y(t)$ is related to the three inputs as follows:

$$Y(s) = F(s)\left(Y^0(s) - N(s)\right) + S(s)D(s) + V(s)W(s)$$

hence, in order to make $y(t) \approx y^0(t)$, we need $|F(j\omega)| \approx 1$ in the bandwidth of the reference signal $y^0(t)$ and $|F(j\omega)| \approx 0$ in the bandwidth of the measurement noise $n(t)$. Notice that this is possible provided that $\left|Y^0(j\omega)\right| \approx 0, \omega \gg \omega_c$ and $|N(j\omega)| \approx 0, \omega \ll \omega_c$, that is the crossover frequency ω_c is selected to well separate the bandwidths of the reference $y^0(t)$ and the one of the measurement noise $n(t)$. Moreover, we also need $|S(j\omega)| \approx 0$ in the bandwidth of the disturbance $d(t)$, as well as $|V(j\omega)| \approx 0$ in the one of the load disturbance $w(t)$.
From (7.2), (7.7), and (7.10), all these requirements are achieved by making **the magnitude of $L(j\omega)$ sufficiently large for frequencies that are lower than the crossover, i.e. $\forall\omega \ll \omega_c$ and sufficiently small for higher frequencies, that is $\forall\omega \gg \omega_c$.**

The error signal $e(t)$ is related to the three inputs as follows:

$$E(s) = S(s)\left(Y^0(s) - D(s)\right) - F(s)N(s) - V(s)W(s)$$

hence, in order to make $e(t) \approx 0$, we need $|S(j\omega)| \approx 0$ in the bandwidth of both the reference signal $y^0(t)$ and the disturbance $d(t)$. Notice that this is possible provided that both $\left|Y^0(j\omega)\right| \approx 0, \omega \gg \omega_c$ and $|D(j\omega)| \approx 0, \omega \gg \omega_c$, that is the crossover frequency ω_c is selected to include both the bandwidth of the reference $y^0(t)$ and the one of the disturbance $d(t)$. Moreover, we also need $|F(j\omega)| \approx 0$ in the bandwidth of the measurement noise $n(t)$, as well as $|V(j\omega)| \approx 0$ in the one of the load disturbance $w(t)$.
From (7.7), (7.2), and (7.10), all these requirements are again achieved by making **the magnitude of $L(j\omega)$ sufficiently large for frequencies that are lower than the crossover, i.e. $\forall\omega \ll \omega_c$ and sufficiently small for higher frequencies, i.e. $\forall\omega \gg \omega_c$.**

Finally, the control signal $u(t)$ is related to the three inputs as follows:

$$U(s) = Q(s)\left(Y^0(s) - D(s) - N(s)\right) - F(s)W(s)$$

hence, in order to moderate the effort of the actuator, i.e. $u(t) \approx 0$, we need $|Q(j\omega)| \approx 0$ in the bandwidth of both the reference signal $y^0(t)$, the disturbance $d(t)$, and the measurement noise $n(t)$. Notice from (7.9) this is just possible for the measurement noise $n(t)$. Moreover, we also need $|F(j\omega)| \approx 0$ in the

bandwidth of the load disturbance $w(t)$, which is also not possible.
From (7.9) and (7.2), all the possible requirements are again achieved by making **the magnitude of $L(j\omega)$ sufficiently large for frequencies that are lower than the crossover, i.e. $\forall\omega \ll \omega_c$ and sufficiently small for higher frequencies, i.e. $\forall\omega \gg \omega_c$.**

Chapter 8

LTI SISO control systems synthesis

In Chapter 7, we have seen how the characteristics of the frequency response of the loop transfer function $L(s)$ are influencing the behaviour of the closed-loop system. In this Chapter, we rather look at the opposite problem: we will try to shape the frequency response of the loop transfer function based on some prescribed requirements we want the closed-loop system to have.

8.1 Handling of the requirements

In the light of the analyses reported in Chapter 7, it should be clear that **all requirements for the closed-loop system** can be translated into equivalent **constraints on the frequency response of the loop transfer function** $L(s)$. In the following, we will address how this process can be handled in a systematic way. At the end of this process, i.e. when all requirements will be somehow referred to the loop transfer function, a candidate form for $L(s)$ will be easily derived, from which the controller

$$R(s) = \frac{L(s)}{G(s)}$$

will immediately follow.

8.1.1 Stability requirement

In this Section, the stability requirement for the closed-loop system will be revised from another perspective. We actually want to investigate how the frequency response of the loop transfer function $L(s)$ is expected to be shaped in order to guarantee the asymptotic stability for the closed-loop system, i.e. a positive phase margin $\psi_m > 0$.

Example 48. *Consider the following loop transfer function*

$$L_1(s) = \frac{10^{6/20}}{(1+s)^2}$$

which is assumed to be obtained without cancellations within the right half-plane. The corresponding Bode's diagrams are shown in Fig. 8.1 from which $\omega_c = 1$ *rad/s and* $\psi_m = 90$ deg. *Notice that the phase margin can be obtained as*

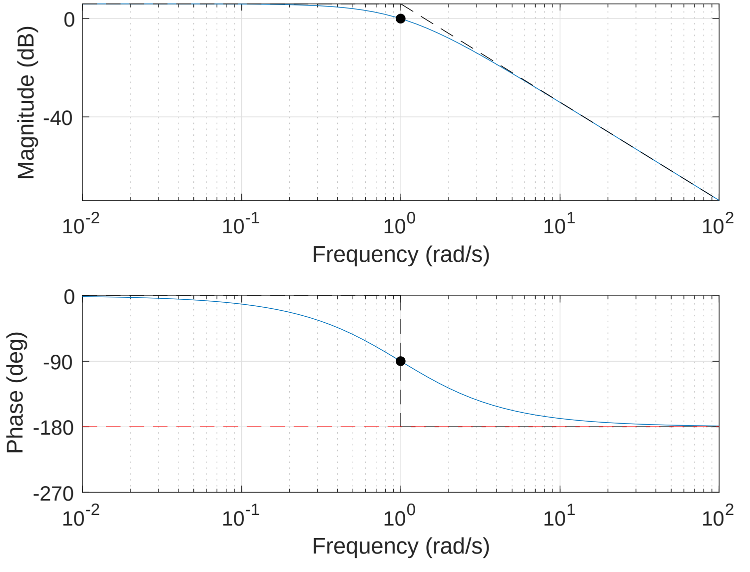


Figure 8.1: Bode's diagrams of the frequency response of $L_1(s)$.

$$\psi_m = \pi - 2\text{atan}(\omega_c)$$

which is quite sensitive to variation in the crossover frequency from the actual value $\omega_c = 1$, *in fact*

$$\frac{\partial \psi_m}{\partial \omega_c} = -\frac{2}{1+\omega_c^2} = -1$$

Consider now an alternative loop transfer function

$$L_2(s) = \frac{10}{1+10s}$$

The corresponding Bode's diagrams are shown in Fig. 8.2 from which $\omega_c = 1$ *rad/s and* $\psi_m \approx 90$ deg, *hence approximately the same of those obtained for* $L_1(s)$. *Notice that the phase margin can be obtained as*

$$\psi_m = \pi - \text{atan}(10\omega_c)$$

which is less sensitive to possible variations in the crossover frequency from the value $\omega_c \approx 1$, *in fact*

$$\frac{\partial \psi_m}{\partial \omega_c} = -\frac{10}{1+100\omega_c} \approx -0.1$$

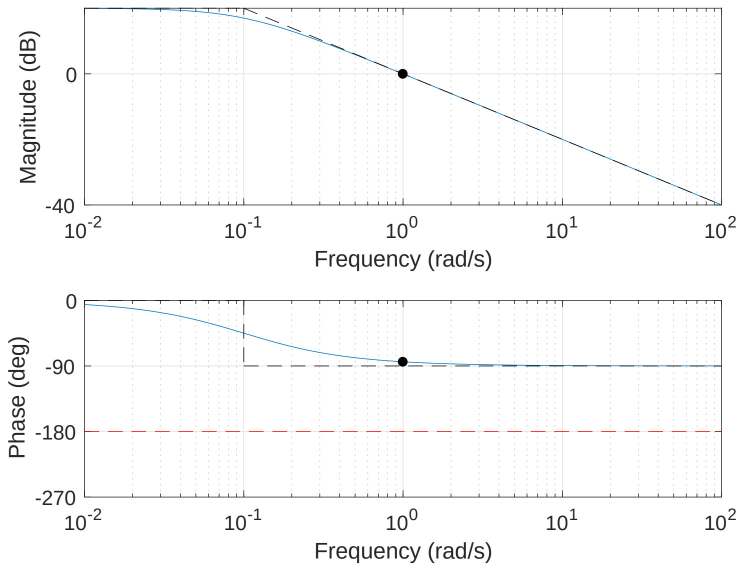


Figure 8.2: Bode's diagrams of the frequency response of $L_2(s)$.

From the previous example, we have seen that the presence of singularities (poles or zeros) near the crossover frequency ω_c can severely affect the sensitivity of the phase margin with respect to variations in the crossover frequency around its nominal value. The possible uncertainties in the evaluation of the crossover frequency ω_c should not be surprising. Its value, in fact, depends on the magnitude of the frequency response of the loop transfer function $L(s)$ which, in turn, depends on the magnitude of the frequency response of the system to be controlled for which $G(s)$ only represents a (reasonably good, yet not perfect) model.
As a rule of thumb, the crossover frequency should be far enough from singularities. This way, in the neighbourhood of ω_c, **the behaviour of the phase diagram should be substantially flat**, so that the corresponding evaluation of the phase margin ψ_m will result more robust with respect to uncertainties.

Beside requiring the magnitude diagram of the frequency response of the loop transfer function $L(s)$ to have no singularities in the neighbourhood of the crossover frequency ω_c, it is interesting to understand the way this crossing should happen. From the discussions in Chapter 4, we have seen that the slope of the magnitude diagram at any frequency ω is equal to the difference between zeros and poles (accounting for their multiplicity) appearing before ω. Therefore, the slope of the magnitude diagram of $L(j\omega)$ at the crossover frequency ω_c is equal to the difference between the number of zeros and the number of poles before the crossover frequency. As the crossing should happen from top to bottom, the slope of the magnitude diagram has to be negative, meaning that the loop transfer function $L(s)$ should have more poles than zeros before ω_c. From the discussion above, these singularities should all happen well before ω_c. Under these circumstances, since the phase shift associated to each singularity with time constant $T : |T| \gg 1/\omega_c$ is

$$\angle (1 + j\omega_c T)^{\pm 1} = \pm \mathrm{atan}(T\omega_c) \approx \pm \mathrm{sign}(T) \frac{\pi}{2}$$

the sole possibility to have a negative slope at the crossover frequency and a positive phase margin (i.e. $\angle L(j\omega_c) > -\pi$) is to cross the 0 dB axis with slope -20 dB/decade, i.e. to have **exactly one more pole than zeros before** ω_c.

8.1.2 Steady-state requirements

In many practical applications, the precision of the control system in tracking the reference signal $y^0(t)$ is of paramount importance. In the motivating example introduced in Chapter 1, a significant steady-state error in the cruise control might compromise the integrity of the vehicle, the safety of the driver, and might cause the driver to receive a fine. In the last Section of Chapter 7, we have seen that the relationship between the error signal and the inputs of the closed-loop system can be written as

$$E(s) = S(s)\left(Y^0(s) - D(s)\right) - V(s)W(s)$$

where we already assumed the measurement noise $n(t)$ to have null components at low frequencies. If the controller to be designed $R(s)$ will be able to make the closed-loop system asymptotically stable, the final value theorem asserts that

$$e_\infty = \lim_{t \to +\infty} e(t) = \lim_{s \to 0} sE(s)$$

By substitution, we obtain

$$e_\infty = \lim_{s \to 0} sS(s)\left(Y^0(s) - D(s)\right) - \lim_{s \to 0} sV(s)W(s)$$

Since

$$\lim_{s \to 0} R(s) - \frac{\mu_R}{s^{g_R}} = 0, \lim_{s \to 0} G(s) - \frac{\mu_G}{s^{g_G}} = 0$$

we can further write

$$e_\infty = \lim_{s \to 0} \frac{s^{g_L+1}}{s^{g_L} + \mu_L}\left(Y^0(s) - D(s)\right) - \lim_{s \to 0} \frac{\mu_G s^{g_R+1}}{s^{g_L} + \mu_L} W(s) \tag{8.1}$$

Therefore, given a reference signal $y^0(t)$, the disturbance $d(t)$ and the load disturbance $w(t)$ that typically characterises certain applications, the requirement on the steady-state error e_∞ (for example $e_\infty = 0$ or $|e_\infty| \leq e_\infty^{\max}$) can be obtained by **properly selecting the generalised gain μ_R and the type g_R of the controller**, possibly based on the generalised gain and the type of the system under control, i.e. μ_G and g_G, respectively, and therefore the corresponding ones for the loop transfer function $L(s)$.

Example 49. *[Handling the steady-state requirement] Consider the system described by the following transfer function*

$$G(s) = \frac{1-s}{s}$$

and assume the requirement on the state-state condition of the error to be

$$|e_\infty| \leq 0.01$$

when $y^0(t) = \bar{y}^0 t, t \geq 0$ with $\left|\bar{y}^0\right| \leq 1$ and $w(t) = \bar{w}t^2, t \geq 0$ with $|\bar{w}| \leq 0.1$.
Assuming the closed-loop system to be asymptotically stable, the requirements on the generalised gain and type of the controller can be obtained by applying (8.1) to the case at hand, with

$$Y^0(s) = \frac{\bar{y}^0}{s}, W(s) = \frac{\bar{w}}{s^2}, D(s) = 0$$

and therefore, knowing that $\mu_G = 1$ and $g_G = 1$, one obtains:

$$|e_\infty| \leq \left|\lim_{s\to 0} \frac{s^{g_R+1}}{s^{g_R+1}+\mu_R}\right| + 0.1\left|\lim_{s\to 0} \frac{s^{g_R-1}}{s^{g_R+1}+\mu_R}\right|$$

The first term is null $\forall g_R \geq 0$ and $\forall \mu_R > 0$, whilst for the second we have two possibilities: either $g_R = 1, \mu_R \geq 10$ or $g_R \geq 2, \forall \mu_R > 0$, or equivalently

$$g_L = 2, \mu_L \geq 10 \text{ or } \mu_L \geq 3, \forall \mu_L > 0$$

8.1.3 Requirements on transients

The duration of transients mainly depends on the position of closed-loop poles which, in turn, depends on the crossover frequency ω_c. In particular, as we have seen in Chapter 7, the complementary sensitivity $F(s)$, can be approximated with a first-order transfer function, i.e.

$$F(s) \approx \frac{1}{1+s/\omega_c}$$

for sufficiently high phase margins, e.g. $\psi_m > 80$ deg, or with a second-order one

$$F(s) \approx \frac{\omega_c^2}{s^2+2\xi\omega_c s+\omega_c^2}, \xi = \sin\left(\frac{\psi_m}{2}\right)$$

whenever $\psi_m < 40$ deg. Then, assuming one is interested in analysing the transient of $y(t)$ in response to a step reference of $y^0(t)$, according to the analysis of first- and second-order step responses in Chapter 3, the corresponding 1%-settling time would be

$$T_{s1\%} \approx \frac{\log_e(100)}{\omega_c} \approx \frac{4.6}{\omega_c}$$

for high values of the phase margin ψ_m, or

$$T_{s1\%} \approx \frac{\log_e(100)}{\xi\omega_c} \approx \frac{4.6}{\xi\omega_c}$$

for low values.

In one case, the requirement on transients can be addressed by a proper constraint on the crossover frequency ω_c, provided that it is possible to obtain a reasonably high phase margin. Whenever this is not possible, the requirement on the settling time has to be addressed by constraining both the crossover frequency ω_c and the phase margin ψ_m.

Example 50. *[Handling requirements on transients] Assume the controller has to be designed so that for a step reference $y^0(t)$, the output $y(t)$ settles to its steady-state value in no more than 3 seconds with a maximum overshoot of 40%.*
Whenever it is possible to obtain a sufficiently large phase margin (e.g. $\psi_m > 80$ deg) to justify a first-order approximation of the complementary sensitivity (hence with no overshoot), it is just necessary to require $\omega_c \geq 4.6/3 = 1.5317$ rad/s. Otherwise, we have seen in Chapter 3 that for a second-order response, the maximum overshoot is related to the damping as

$$S_{\%} = 100e^{-\frac{\pi\xi}{\sqrt{1-\xi^2}}}$$

Hence, imposing

$$100e^{-\frac{\xi\pi}{\sqrt{1-\xi^2}}} \leq 40$$

one obtains a requirement on the minimum damping, that is $\xi \geq 0.28$, which corresponds to $\psi_m \geq 2\arcsin(\xi) \approx 32.5$ deg. Therefore, one needs $\omega_c \geq 4.6/(3\xi) = 5.4704$ rad/s.
In summary, in order to meet the given requirement on the duration of transients, one has to satisfy either $\omega_c \geq 1.5317, \psi_m > 80$ deg, or, alternatively, the less demanding requirement $\omega_c \geq 5.4704, \psi_m > 32.5$ deg.

8.1.4 Attenuation requirements

Attenuation requirements are handled using the frequency response theorem and therefore will be satisfied only after the transients. Typical requirements are specified through the relationship between the amplitude of harmonics of one of the input, i.e. $d(t)$, $w(t)$ or $n(t)$, and the corresponding amplitude of the corresponding harmonic response of one output, i.e. $y(t)$, $y(t)$, or $u(t)$.
The key idea to handle these kind of requirement is to consider which transfer function in the "Gang of four" is related to the given input/output pair, and to use one of the approximations derived in Chapter 7 to translate the requirement in terms of constraints on the magnitude of the frequency response of the loop transfer function $L(s)$.

Example 51. *[Handling attenuation requirements] Given the system*

$$G(s) = \frac{10}{1+10s}$$

assume for one wants to attenuate on $y(t)$ *the effects of* $n(t) = \bar{n}\sin(\bar{\omega})$ *of at least* 30 dB *for* $\bar{\omega} > 100$ rad/s*, as well as of at least* 20 dB *the effects of* $w(t) = \bar{w}\sin(\tilde{\omega}t), \tilde{\omega} < 0.3$ rad/s *on the output* $y(t)$*.*

To handle the former, one has to consider the transfer function from the measurement noise $n(t)$ *to the controlled variable* $y(t)$ *that is the complementary sensitivity*

$$F(s) = \frac{L(s)}{1+L(s)}$$

The sole possibility to handle this requirement is to have $\omega_c < 100$ rad/s*. Correspondingly, one can use the high frequency approximation of the magnitude of the frequency response of* $F(s)$*, that is*

$$|F(j\omega)| \approx |L(j\omega)|$$

Then, assuming the closed-loop system to eventually result asymptotically stable, one can apply the frequency response theorem and require

$$|L(j\omega)| < 10^{-30/20} \approx 0.0316, \forall\omega > 100$$

Hence, the magnitude diagram of $L(j\omega)$ *has to stay below the one of* $10^{(-30/20)}$ *for* $\omega > 100$ rad/s*.*

To handle the latter requirement, one has to focus on the transfer function from the load disturbance $w(t)$ *to the output variable* $y(t)$*, i.e., apart from the sign, the output sensitivity* $V(s)$*. To handle this requirement we need to have* $\omega_c > 0.3$ rad/s*. Using the corresponding approximation of the output sensitivity, one is actually requiring*

$$\frac{|G(j\omega)|}{|L(j\omega)|} < 10^{-20/20} = 0.1, \forall\omega < 0.3$$

or equivalently

$$|L(j\omega)| > 10\,|G(j\omega)|\,, \forall\omega < 0.3$$

Hence, the magnitude diagram of $L(j\omega)$ *has to stay above the one of* $10\,|G(j\omega)|$ *for* $\omega < 0.3$ rad/s*.*

Figure 8.3 reports the graphical representation of the two attenuation requirements and the magnitude of the frequency response of a possible loop transfer function

$$L(s) = \sqrt{10}\frac{1+s}{s^2}$$

which satisfies both of them with $\omega_c = 3.3$ rad/s *and* $\psi_m = 73.2$ deg.

Notice that, given the loop transfer function $L(s)$ *and the transfer function of the system* $G(s)$*, one can easily derive the transfer function of the controller as*

$$R(s) = \frac{L(s)}{G(s)} = \frac{1}{\sqrt{10}}\frac{(1+s)(1+10s)}{s^2}$$

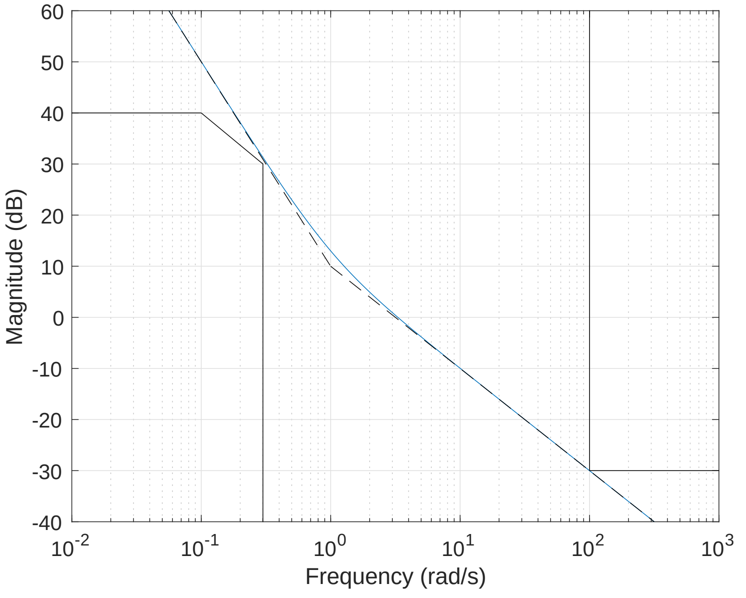


Figure 8.3: Graphical rendering of the two attenuation requirements of this example and the magnitude diagram of the frequency response of a candidate loop transfer function $L(s)$.

8.2 Synthesis of the controller

In the previous Sections, we have seen that all the requirements for the closed-loop system can be translated in terms of properties of the loop transfer function. For example, the requirement on closed-loop stability has been translated into the shape of the loop transfer function and $\psi_m > 0$, the steady-state requirements have been translated into constraints on the generalised gain μ_L and type g_L of $L(s)$, etc. Table 8.1 summarises all the findings.

Once the loop transfer function $L(s)$ has been designed to satisfy all these requirements, the controller (s) can be simply computed as

$$R(s) = \frac{L(s)}{G(s)}$$

We are now ready to present a complete example where the transfer function of the controller $R(s)$ is designed based on a set of requirements for the closed-loop system.

Example 52. *[Mass with friction - a complete example] Consider the mass with friction of Fig. 6.1 with $m = 50$ and $r = 3$, and assume a closed-loop controller $R(s)$ has to be designed to obtain the following requirements for the closed-loop system:*

Closed-loop requirement	**Constraint on $L(s)$**
asymptotic stability	crossing 0 dB with slope $-20dB/decade$ well far from singularities, $\psi_m > 0$
steady-state error	generalised gain and type
attenuation of disturbances on output	magnitude sufficiently large in the corresponding bandwidth
attenuation of noise on output	magnitude sufficiently small in the corresponding bandwidth
reduced settling time	crossover frequency sufficiently high
reduced overshoot	phase margin sufficiently high

Table 8.1: Typical requirements for the closed-loop system and their translation in terms of properties of the loop transfer function $L(s)$.

1. *asymptotic stability;*
2. *null steady-state error for $v^0(t) = \bar{v}, t \geq 0$;*
3. *absolute value of the steady-state error lower than 0.1 when $w(t) = t, t \geq 0$;*
4. *effects of measurement noise $n(t)$ attenuated of at least 3 dB on the control variable $f(t)$ for $\omega > 100\ rad/s$;*
5. *a settling time of the response of the output $v(t)$ for a step reference $v^0(t)$ no higher than 12 seconds.*

To handle the first requirement, we will just need to enforce the applicability of the Bode's criterion and check that $\psi_m > 0$.

As per the second requirement, we need to consider the transfer function from the reference $v^0(t)$ to the error $e(t)$, i.e. the sensitivity transfer function, and require

$$e_\infty = \lim_{s\to 0} sE(s) = \bar{v}\lim_{s\to 0} S(s) = \bar{v}\lim_{s\to 0}\frac{1}{1+\dfrac{\mu_L}{s^{g_L}}} = 0$$

which is satisfied for $g_L \geq 1$ and $\forall \mu_L > 0$.

The third requirement can be handled similarly to the previous one by considering the transfer function from the load disturbance $w(t)$ to the error $e(t)$ that is, apart from the sign, the output sensitivity

$$|e_\infty| = \left|\lim_{s\to 0} sE(s)\right| = \left|\lim_{s\to 0} sV(s)\frac{1}{s^2}\right| = \left|\lim_{s\to 0}\frac{\dfrac{s^{-1}}{r}}{1+\dfrac{\mu_L}{s^{g_L}}}\right| = \left|\lim_{s\to 0}\frac{\dfrac{s^{g_L-1}}{r}}{s^{g_L}+\mu_L}\right| \leq 0.1$$

which holds for $g_L \geq 2$ and $\forall\mu_L$ or $g_L = 1$ and $\mu_L \geq 10/r$.

As per the fourth requirement, we can consider the frequency response theorem as applied to the transfer function between the measurement noise $n(t)$ and the control input $f(t)$ which is, apart from the sign, the control sensitivity $Q(s)$. For high frequencies, the control sensitivity can be approximated with the transfer function of the controller $R(s)$ and therefore we have to require

$$|R(j\omega)| \leq 10^{-3/20} = 0.7079, \forall\omega > 100$$

Moreover, since $R(s) = L(s)/G(s)$ *we can equivalently require*

$$|L(j\omega)| \leq 0.7079\,|G(j\omega)|\,, \forall \omega > 100\ rad/s$$

Finally, assuming the resulting phase margin to be sufficiently high (e.g. $\psi_m > 80$ deg*), the last requirement on the settling time can be simply translated as follows*

$$\omega_c \geq \frac{4.6}{12} = 0.3829\ rad/s$$

If this constraint will result impossible to be satisfied, we will use the second-order approximation of the complementary sensitivity, instead.

The graphical representation of all the derived requirements is reported in Fig. 8.4. The next step is then to draw the magnitude diagram of $L(j\omega)$ *and finally check if the phase margin* ψ_m *will result sufficiently high to justify the hypothesis.*

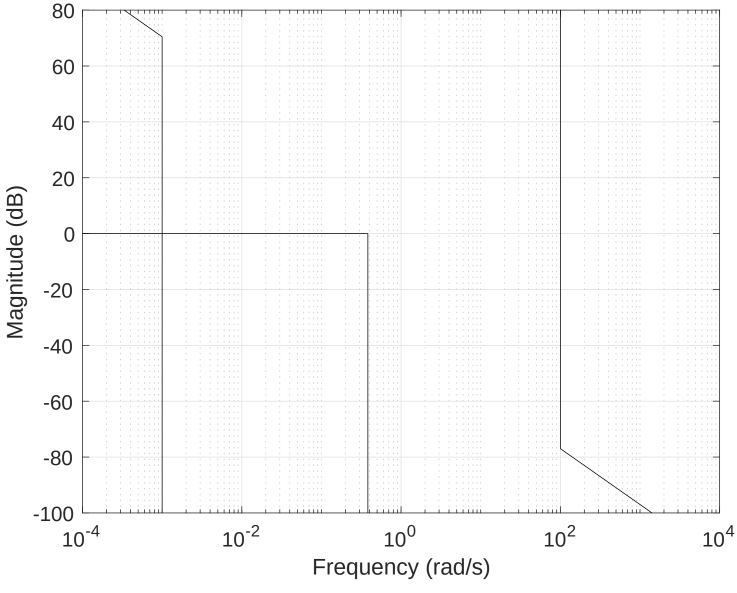


Figure 8.4: Graphical rendering of the requirements of this example.

As a first attempt in designing the magnitude diagram of $L(j\omega)$ *we can focus on*

$$L(s) = \frac{0.4}{s}$$

which has a crossover frequency of $\omega_c = 0.4\ rad/s$ *and a phase margin of* $\psi_m = 90$ deg *and hence satisfy both the first and last requirements. The corresponding magnitude diagram is shown in Fig. 8.5.*

As one can see from Fig. 8.5, the other three requirements are not yet satisfied. Then, we can set $g_L = 2$ *and add a zero at least one decade before* $\omega_c = 0.4\ rad/s$*. The corresponding loop transfer function would then result*

$$L(s) = \frac{0.4}{100}\frac{1 + 100s}{s^2}$$

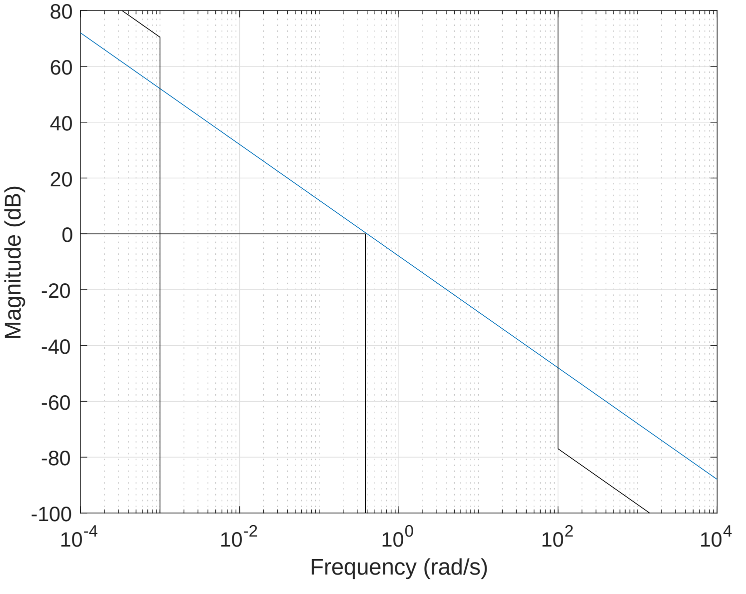


Figure 8.5: First attempt in designing $L(s)$.

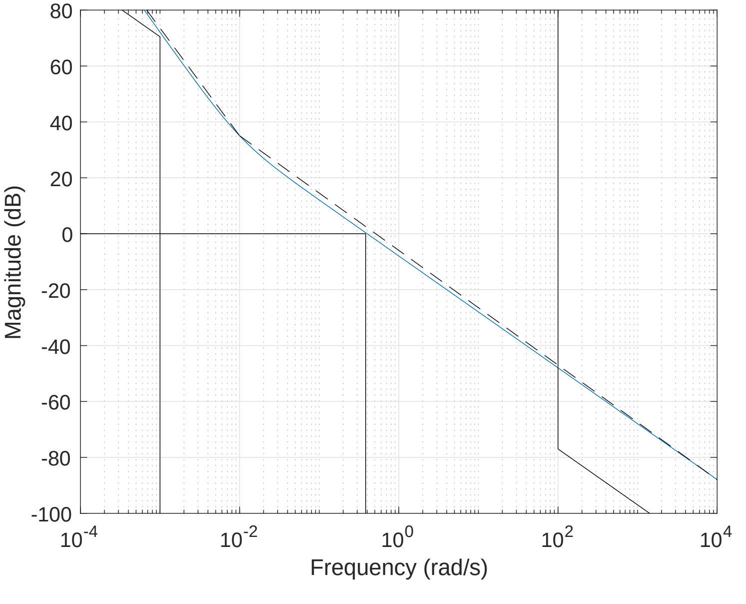


Figure 8.6: Second attempt in designing $L(s)$.

which corresponds to a crossover frequency $\omega_c = 0.4$ *rad/s and a phase margin of* $\psi_m = 88.6$ deg *and to the magnitude diagram in Fig. 8.6.*

The first three requirements are satisfied as well as the last one. However, this solution is not yet consistent with the fourth requirement. Adding a pole approximately one decade after ω_c would produce a loop transfer function $L(s)$ satisfying all the constraints. Consider then

$$L(s) = \frac{0.4}{100} \frac{1+100s}{s^2(1+s/3.5)}$$

Correspondingly, we obtained $\omega_c = 0.398$ rad/s and $\psi_m = 82.1$ deg, while the magnitude diagram of $L(j\omega)$ is reported in Fig. 8.7.

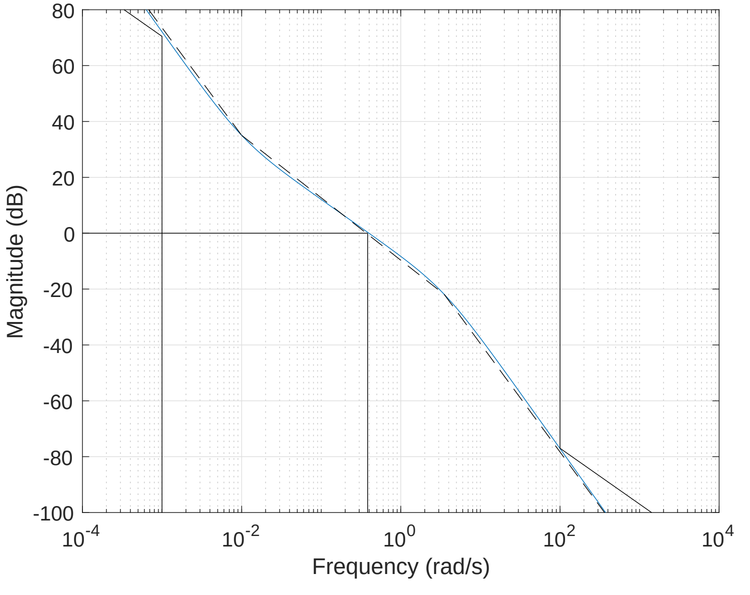


Figure 8.7: Final attempt in designing $L(s)$.

As one can notice, the proposed loop transfer function now satisfies all the prescribed requirements. The transfer function of the corresponding controller $R(s)$, i.e. the output of the design process, can be simply obtained as

$$R(s) = \frac{L(s)}{G(s)} = \frac{\dfrac{0.4}{100}\dfrac{1+100s}{s^2(1+s/3.5)}}{\dfrac{1}{50s+3}} = \frac{0.4}{100}\frac{(1+100s)(50s+3)}{s^2(1+s/3.5)}$$

As a final verification, Figure 8.8 reports the step response of the output velocity $v(t)$, the response of the error $e(t)$ for a load disturbance $w(t) = 0.1t, t \geq 0$ and the behaviour of the control signal $f(t)$ when the measurement noise is $n(t) = 0.1\sin(100t)$. From Fig. 8.8 notice that:

1. *the steady state error is null after a quite long transient both in case of a step reference and a linearly increasing load disturbance;*
2. *the response of the output is quite fast to get close to the reference value;*
3. *the amplitude of the oscillations in the control signal is less than approximately* 0.07 N *when the measurement noise has amplitude* 0.1 m/s, *hence the attenuation effect is of at least* 1.4 *times (3 dB).*

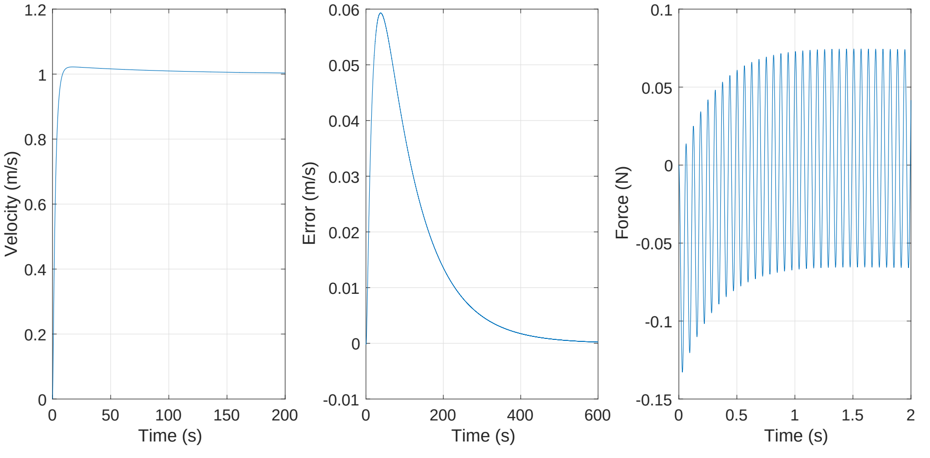


Figure 8.8: Validation scenarios. From left to right: step response of $v(t)$, response of $e(t)$ for $w(t) = t, t \geq 0$, and behaviour of $f(t)$ for $n(t) = 0.1\sin(100t)$.

8.3 Design limitations

In the example of the previous Section, we have seen that apart from the requirement on the attenuation of the effects of the measurement noise $n(t)$, where was no particular upper bound on the selection of the crossover frequency ω_c. Instead, a lower bound was introduced to handle the requirements on the transients. In this Section, we will investigate two situations characterised by a limit on the maximum achievable crossover frequency ω_c.

8.3.1 Delayed actuation or measurement

Consider now a LTI SISO system to be controlled characterised by the following transfer function

$$G(s) = \tilde{G}(s)\, e^{-s\tau}$$

where $\tilde{G}(s)$ represents a generic rational transfer function (ratio between two polynomials) with all singularities (zeros and poles) within the left half-plane, while $e^{-s\tau}, \tau \geq 0$ represents a delay in either the actuation or the measurement. For a generic controller $R(s)$, the corresponding loop transfer function results as follows

$$L(s) = R(s)\, G(s)\, e^{-s\tau} = \tilde{L}(s)\, e^{-s\tau}$$

with $\tilde{L}(s) = R(s)\,\tilde{G}(s)$. As already discussed in Chapter 4, the presence of the delay does not affect the magnitude diagram of the frequency response of $L(s)$, i.e.

$$\forall\omega : |L(j\omega)| = \left|\tilde{L}(j\omega)\right| \left|e^{-j\omega\tau}\right| = \left|\tilde{L}(j\omega)\right|$$

hence the crossover frequency ω_c is not affected by the possible presence of the delay.

As discussed in this Chapter, in order to cross the 0 dB axis with slope -20 $dB/$decade, we expect one more pole than zeros before the crossover frequency. Moreover, all singularities (zeros and poles) should be placed well far away from the crossover frequency ω_c, hence

- the phase shift associated to each left half-plane singularity **after** the crossover frequency, i.e. with time constants $T \ll 1/\omega_c, T > 0$ is negligible, in fact
$$\angle(1 + j\omega_c T)^{\pm 1} = \pm\text{atan}(T\omega_c) \approx 0$$
- the phase shift due to left half-plane singularities **before** the crossover frequency, i.e. $T \gg 1/\omega_c, T > 0$, is
$$\angle(1 + j\omega_c T)^{\pm 1} = \pm\text{atan}(T\omega_c) \approx \pm\text{sign}(T)\frac{\pi}{2} = \pm\frac{\pi}{2}$$

Since we expect one more pole than zeros before the crossover frequency, we have

$$\angle\tilde{L}(j\omega_c) = -\frac{\pi}{2}$$

and therefore

$$\angle L(j\omega_c) = \angle\tilde{L}(j\omega_c) - \omega_c\tau = -\frac{\pi}{2} - \omega_c\tau$$

Correspondingly, the phase margin results as follows

$$\psi_m = \pi - |\angle L(j\omega_c)| = \pi - \left(\frac{\pi}{2} + \omega_c\tau\right) = \frac{\pi}{2} - \omega_c\tau$$

For the asymptotic stability of the closed-loop system, a positive phase margin is required, therefore

$$\psi_m = \frac{\pi}{2} - \omega_c\tau > 0$$

As a conclusion, the presence of a delay $\tau > 0$ in the system to be controlled (either in the actuator or in the sensor) limits from above the value of the crossover frequency ω_c:

$$\omega_c < \frac{\pi}{2\tau} \tag{8.2}$$

Example 53. *[Delay and upper-bound on the crossover frequency] Consider the following loop transfer function*

$$L(s) = k\frac{1+s}{s^2}e^{-0.1s}, k > 0$$

For sufficiently high frequencies, the magnitude diagram of $L(j\omega)$ can be approximated with the one of $k/j\omega$ and therefore $\omega_c = k$. Using the estimate of the maximum crossover frequency in (8.2), we can expect the closed-loop system to be asymptotically stable for $k < \pi/0.2 \approx 15.7080$ rad/s. Figure 8.9 shows the Bode's diagrams of $L(j\omega)$ for unit gain $k = 1$.

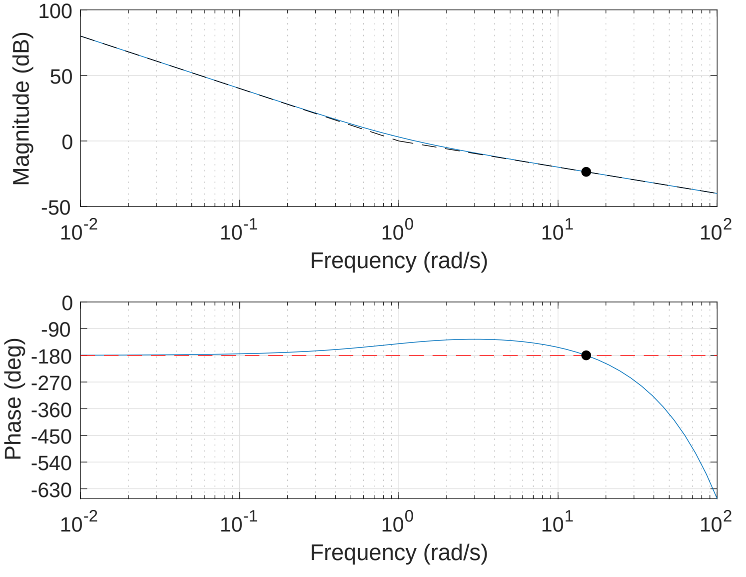


Figure 8.9: Bode's diagrams of the frequency response of the loop transfer function of this example for $k = 1$.

As one can see from the Bode's diagrams, the phase margin ψ_m vanishes for $\omega_c = 15.0421$ rad/s, which is very close to the value of 15.7080 rad/s one would have predicted using (8.2).

8.3.2 Right half-plane zero

Consider now a LTI SISO system to be controlled characterised by the following transfer function

$$G(s) = \tilde{G}(s)(1 - s\tau)$$

where $\tilde{G}(s)$ represents a strictly proper transfer function with no singularities in the right half-plane and $\tau > 0$ is the time constant of the unique zero in the right half-plane. For a generic controller $R(s)$, the corresponding loop transfer function results as follows

$$L(s) = R(s)G(s)(1 - s\tau) = \tilde{L}(s)(1 - s\tau)$$

with $\tilde{L}(s) = R(s)\tilde{G}(s)$. Notice that, differently from the case of the delay, this time, the presence of the right half-plane zero does affect the magnitude diagram of the frequency response of $L(s)$. It is worth to remember that, as discussed in Chapter 6, it is not possible to ensure closed-loop asymptotic stability by designing a controller that cancels out the presence of the right half-plane zero in $G(s)$ with a corresponding pole.

Again, in order to cross the 0 dB axis with slope -20 dB/decade, we expect in the loop transfer function $L(s)$ one more pole than zeros before the crossover frequency ω_c. In addition, as already discussed, all singularities (zeros and poles) are to be placed well far away from the crossover frequency so that the phase shift associated to zeros and poles **after** the crossover frequency, i.e. having time constant $T \ll 1/\omega_c, T > 0$ is negligible. Differently from the previous case, there are two options for the position of the right half-plane zero: before and after the crossover frequency ω_c.

Right half-plane zero before ω_c

If crossover frequency ω_c is set at a higher frequency with respect to the one of the right half-plane zero, i.e. $\omega_c > 1/\tau$, in order to cross the 0 dB axis with slope -20 dB/decade, two more poles than zeros are expected before ω_c, in this case

$$\angle L(j\omega_c) = \angle\tilde{L}(j\omega_c) - \text{atan}(\tau\omega_c) = -\pi - \text{atan}(\tau\omega_c)$$

and the corresponding phase margin

$$\psi_m = \pi - |-\pi - \text{atan}(\tau\omega_c)| = -\text{atan}(\tau\omega_c) < 0$$

cannot be positive.

Right half-plane zero after ω_c

In turn, if the crossover frequency is selected to be lower than the frequency of the right half-plane zero, i.e. $\omega_c < 1/\tau$, only one more pole than zeros is expected in order to cross 0 dB axis at ω_c with slope -20 dB/decade. Correspondingly

$$\angle L(j\omega_c) = \angle\tilde{L}(j\omega_c) - \text{atan}(\tau\omega_c) = -\frac{\pi}{2} - \text{atan}(\tau\omega_c)$$

while the phase margin

$$\psi_m = \pi - \left|-\frac{\pi}{2} - \text{atan}(\tau\omega_c)\right| = \frac{\pi}{2} - \text{atan}(\tau\omega_c)$$

can be positive, since $\text{atan}(\tau\omega_c) < \pi/4$ whenever $\omega_c < 1/\tau$.

In summary, the presence in the transfer function $G(s)$ of a right half-plane zero has to be handled by limiting the crossover frequency ω_c from above, i.e. by requiring

$$\omega_c < \frac{1}{\tau} \tag{8.3}$$

Example 54. *[Right half-plane zero and upper-bound on the crossover frequency] Consider the following loop transfer function*

$$L(s) = k\frac{1-s}{s(1+0.1s)}, k > 0$$

Using the estimate of the maximum crossover frequency in (8.3)*, we can expect the closed-loop system to be asymptotically stable whenever* $\omega_c < 1$ *rad/s. Figure 8.10 shows the Bode's diagrams of* $L(j\omega)$ *for unit gain* $k = 1$*. As one can notice from the Bode's diagrams, the*

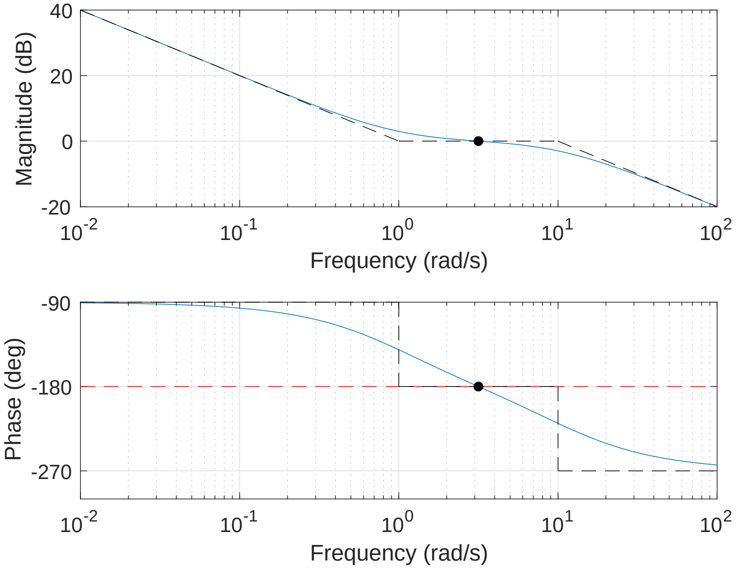


Figure 8.10: Bode's diagrams of the frequency response of the loop transfer function of this example for $k = 1$.

phase margin ψ_m *vanishes for* $\omega_c = 3$ *rad/s. For* $\omega_c = 3$ *rad/s, the crossing of the* 0 *dB axis would happen with a barely horizontal magnitude diagram. On the other hand, the preferred slope of the magnitude diagram of* $L(j\omega)$ *which also guaranteed a positive phase margin is in the bandwidth* $\omega < 1$*, which is exactly the value one would have predicted using* (8.3)*, i.e. the frequency of the right half-plane zero.*

Chapter 9

Feed-forward compensators, PID controllers, and more

In Chapter 8, we have discussed how to design the transfer function $R(s)$ of a feedback controller based on the transfer function $G(s)$ of the system to be controlled and a series of design requirements that the closed-loop system should satisfy. Furthermore, upper bounds on the crossover frequency ω_c were discussed to limit the injection of measurement noise and in the case of non-minimum phased transfer functions $G(s)$. This Chapter will complement the topics of Chapter 8 by introducing typical control schemes.

9.1 Feed-forward compensators

Regulation elements outside the control-loop can be also designed to possibly compensate the limitations in the design of the feedback controller $R(s)$. A feed-forward compensator $C(s)$ is an element placed between the reference signal $y^0(t)$ and the control signal $u(t)$. The corresponding block diagram is reported in Fig. 9.1.

Notice that the presence of transfer function $C(s)$ is only affecting the relationship between the reference $y^0(t)$ and the output variables $e(t), y(t)$, and $u(t)$. In particular, the relationship between the reference $y^0(t)$ and the output $y(t)$ now results as follows

$$Y(s) = \left(F(s) + \frac{C(s)\,G(s)}{1+L(s)} \right) Y^0(s) = \frac{C(s)\,G(s) + L(s)}{1+L(s)} Y^0(s)$$

Hence, if $C(s)$ can be selected so that $C(s)\,G(s) = 1$, then one would obtain the ideal tracking objective

$$Y(s) = Y^0(s)$$

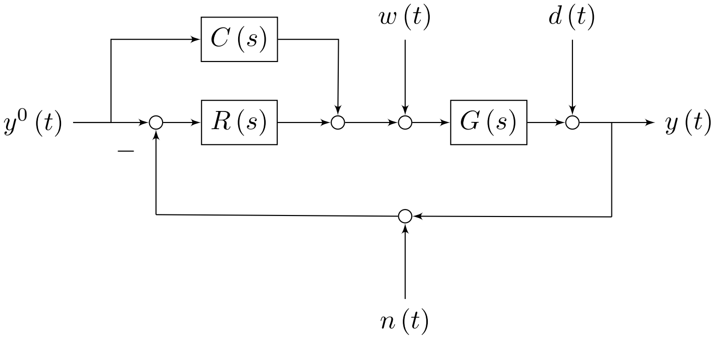


Figure 9.1: A negative feedback control system with feed-forward term $C(s)$.

The ideal tracking objective can be fulfilled if

$$C(s) = \frac{1}{G(s)}$$

represents an **asymptotically stable transfer function**. Requiring $C(s)$ to be asymptotically stable is actually limiting the applicability of this strategy to transfer functions $G(s)$ which are not characterised by right half-plane zeros or delays. More importantly, since most of the times the dynamics of the system is represented by a strictly-proper transfer function $G(s)$, it follows that the ideal $C(s)$ obtained above would not be even a transfer function.

In order to sill benefit of this strategy and overcome its possible limitations, we can design $C(s)$ based on the following less restrictive requirement

$$C(s) : |C(j\omega)| \approx \frac{1}{|G(j\omega)|}, \forall \omega < \omega^{\max}$$

Notice that referring now to the magnitude of $G(j\omega)$ rather than to $G(s)$ will automatically avoid all the problems regarding the possible non-minimum phase components of the system to be controlled. Finally, since

$$|F(j\omega)| \approx 1, \forall \omega \ll \omega_c$$

the design parameter $\omega^{\max}$ should be selected higher than the crossover frequency, i.e. $\omega^{\max} > \omega_c$.

In order to ease the design of transfer function $C(s)$, we introduce the following

Property 15 (Factorisation of a transfer function)**.** *Any SISO transfer function $G(s)$ can be factorised in the product between a transfer function with unit magnitude at all frequencies and a rational transfer function, say $G_{mp}(s)$, with positive generalised gain, no delays, and all singularities (zeros and poles) within the closed left half-plane (minimum phase), i.e.*

$$G(s) = G_{mp}(s)\,\tilde{G}(s)$$

where $\left|\tilde{G}(j\omega)\right| = 1, \forall\omega$, *and therefore* $|G_{mp}(j\omega)| = |G(j\omega)|, \forall\omega$ *(i.e. they have identical magnitude diagrams).*

Example 55. *[Factorisation of a transfer function] Considering the following transfer function*

$$G(s) = -12\frac{1-s}{(1+10s)(1-0.1s)}e^{-s}$$

According to the aforementioned property, transfer function $G(s)$ can be factorised considering

$$G_{mp}(s) = 12\frac{1+s}{(1+10s)(1+0.1s)}$$

and

$$\tilde{G}(s) = -1\frac{(1-s)(1+0.1s)}{(1-0.1s)(1+s)}e^{-s}$$

The verification of the property $\left|\tilde{G}(j\omega)\right| = 1, \forall\omega$ *is left to the reader.*

In the light of the property above, a possible design for the feed-forward compensator can be the following one

$$C(s) = \frac{1}{G_{mp}(s)}$$

which is such that

$$|C(j\omega)| = \frac{1}{|G_{mp}(j\omega)|} = \frac{1}{|G(j\omega)|}$$

Should this still not be a transfer function, high-frequency poles of $G_{mp}(s)$ (i.e. those at frequency $\omega > \omega^{\max}$) can be neglected, or others can be added in $C(s)$ at frequencies higher than $\omega^{\max}$.

Example 56. *[Design of a feed-forward compensator] Consider transfer function*

$$G(s) = 10\frac{1-0.1s}{s(1+s)}$$

As discussed in Chapter 8, the presence of a non-minimum phase zero with time constant $\tau = 0.1$ is practically limiting from above the crossover frequency, i.e.

$$\omega_c < \frac{1}{\tau} = 10\ rad/s$$

Hence, assume that the following controller

$$R(s) = 0.3\frac{1+s}{1+0.1s}$$

has been design to achieve the crossover frequency $\omega_c = 3$ *rad/s and a corresponding phase margin of* $\psi_m = 56.6$ deg.

According to the given guidelines, a starting guess for the compensator $C(s)$ *can be the following one*

$$C_{mp}(s) = 0.1\frac{s(1+s)}{1+0.1s}$$

which is not, however, a transfer function. In turn, by adding a high frequency pole, e.g. at $\omega = 50$ *rad/s, one would obtain*

$$C(s) = 0.1\frac{s(1+s)}{(1+0.1s)(1+0.02s)}$$

The Bode's diagram of the frequency response of $C(s)$ *as compared with those of the ideal compensator* $1/G(s)$ *are shown in Fig. 9.2. As one can see, the designed feed-forward compensator* $C(s)$ *has the same magnitude diagram of the ideal one until at least* $\omega = 10$ *rad/s.*

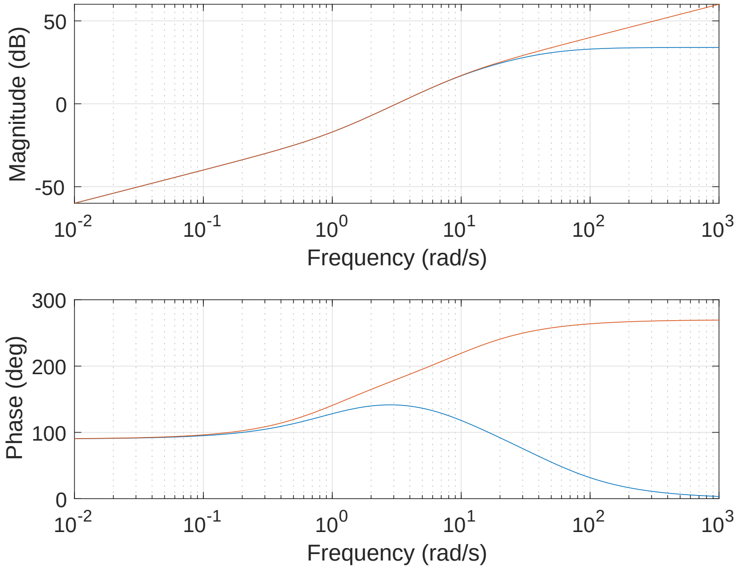


Figure 9.2: Bode's diagrams of the frequency response of the feed-forward compensator $C(s)$ (blue) and the ideal one $1/G(s)$ (red).

Figure 9.3 finally shows the step responses of the output $y(t)$ *with and without the feed-forward compensator* $C(s)$. *Once should notice that while the compensator increases the promptness of the response, it also emphasises both the overshoot and the undershoot.*

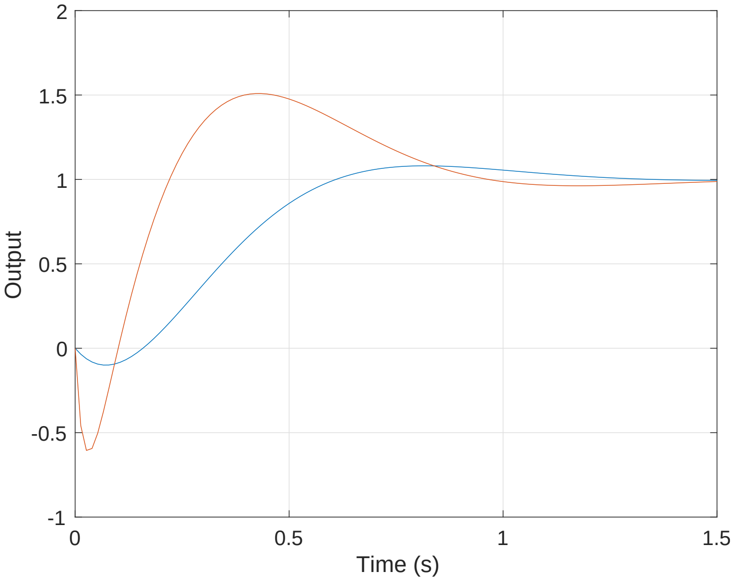


Figure 9.3: Response of $y(t)$ for a step reference $y^0(t)$ with (red) and without (blue) the feed-forward compensator $C(s)$.

9.2 PID controllers

There is a family of controller transfer functions $R(s)$ that is widely adopted in most industrial control problems, thanks to its versatility and the reduced number of parameters to be tuned. PID controllers realise the following control law

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau)\, d\tau + K_d \frac{d}{dt} e(t) \tag{9.1}$$

which is composed of three terms: one proportional (P) to the error $e(t)$[1], one proportional to the integral (I) of the error, the last proportional to the derivative (D) of the error. K_p, K_i and K_d are named the proportional, integral and derivative gains, respectively. Not all the three components are necessarily present, and actually every combination is possible. The proportional gain K_p weights the **current value** of the error, the integral gain K_i relates to the **past values** of the error, while the derivative one K_d tends to anticipate the **future value** of the error.

In the Laplace domain, the relationship between the input of the controller

[1] Actually to the error possibly corrupted by the measurement noise $n(t)$.

Figure 9.4: PID controller (Stanford Research Systems).

$e(t)$ and the control variable $u(t)$ can be written as

$$\begin{aligned} U(s) &= K_p E(s) + \frac{K_i}{s} E(s) + K_d s E(s) = \\ &= \frac{K_i + K_p s + K_d s^2}{s} E(s) = K_p \left(1 + \frac{1}{sT_i} + sT_d\right) E(s) = \\ &= K_p \frac{1 + sT_i + s^2 T_i T_d}{sT_i} E(s) \end{aligned} \tag{9.2}$$

where obviously

$$T_i = \frac{K_p}{K_i}, T_d = \frac{K_d}{K_p}$$

and T_i and T_d are called integral time and derivative time, respectively. Due to the last term, corresponding to the derivative term, the form reported in (9.2) is clearly not a transfer function. Hence, the transfer function of a PID controller is usually obtained by adding a high-frequency pole to the derivative contribution, i.e.

$$\begin{aligned} R(s) = K_p \left(1 + \frac{1}{sT_i} + \frac{sT_d}{1 + \frac{T_d}{N} s}\right) = K_p \frac{sT_i \left(1 + \frac{T_d}{N} s\right) + 1 + \frac{T_d}{N} s + s^2 T_d T_i}{sT_i \left(1 + \frac{T_d}{N} s\right)} = \\ = K_p \frac{1 + s\left(T_i + \frac{T_d}{N}\right) + T_i T_d \left(1 + \frac{1}{N}\right) s^2}{sT_i \left(1 + \frac{T_d}{N} s\right)} \end{aligned} \tag{9.3}$$

Notice that, taking the limit for $N \to +\infty$, the form in (9.3) tends to the one in (9.2). The PID transfer function in (9.3) has generalised gain K_p, a pole in the origin ($s = 0$), a high-frequency pole in $s = -N/T_d$, and, for sufficiently high values of N, two zeros corresponding to the roots of

$$1 + sT_i + s^2 T_i T_d = 0$$

Depending on the values of the integral and the derivative times, i.e. T_i and T_d, zeros can be real or complex and conjugate.

As the range of applications of PID controllers is significantly wide, that International Society of Automation (ISA) has proposed the following standard form

$$U(s) = K_p \left(bY^0(s) - Y(s) + \frac{1}{sT_i} E(s) + \frac{sT_d}{1 + \frac{T_d N}{s}} \left(cY^0(s) - Y(s) \right) \right)$$

Notice that two new parameters, b and c, respectively, have been introduced. The previous relationship can be rewritten, in terms of the Laplace transforms of the error $e(t)$ and the reference $y^0(t)$ as follows

$$U(s) = K_p \left(b - 1 + \frac{sT_d}{1 + \frac{T_d}{N}s} (c-1) \right) Y^0(s) + K_p \left(1 + \frac{1}{sT_i} + \frac{sT_d}{1 + \frac{T_d}{N}s} \right) E(s)$$

in which the first term, say

$$C(s) = K_p \left(b - 1 + \frac{sT_d}{1 + \frac{T_d}{N}s} (c-1) \right) = K_p \frac{(b-1)\left(1 + \frac{T_d}{N}s\right) + (c-1)\, sT_d}{1 + \frac{T_d}{N}s} =$$

$$= K_p \frac{\left((b-1)\frac{T_d}{N} + (c-1)\, T_d \right) s + b - 1}{1 + \frac{T_d}{N}s}$$

represents the transfer function of a feed-forward compensator, while the second term, $R(s)$, is the same of (9.3).

The attentive reader has surely noticed that the real pole of the PID transfer function in (9.3) is actually working as a low-pass filter on the error signal before being differentiated in the derivative term. The presence of measurement noise

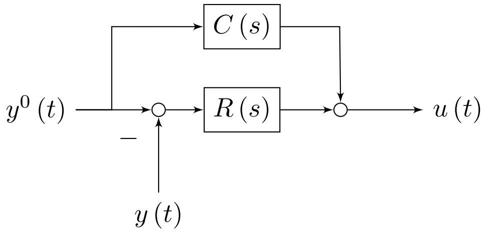


Figure 9.5: Block diagram of the ISA PID.

on the error signal, in fact, usually makes the derivate action quite problematic. Recalling the discussion on the control sensitivity presented in Chapter 7, at high frequencies, i.e. way after the crossover frequency ω_c, we have

$$|Q(j\omega)| \approx |R(j\omega)|$$

Since the control sensitivity represents the relationship between the reference $y^0(t)$ and the control variable $u(t)$, as well as between the measurement noise $n(t)$ and the control variable itself, particular attention should be paid to the high-frequency components of the PID and therefore on the derivate action. In fact, since

$$\lim_{s\to+\infty} R(s) = K_p(1+N)$$

the value of parameter N influences both the amplification of the measurement noise $n(t)$ on the control variable $u(t)$, as well as, for a step response of the reference $y^0(t)$, the corresponding value of $u(0)$, in fact:

$$u(0) = \lim_{s\to+\infty} sU(s) = \lim_{s\to+\infty} Q(s) = \lim_{s\to+\infty} R(s) = K_p(1+N)$$

As visible in the transfer function of the PID controller in (9.3), the tuning of its parameters actually consists in selecting its generalised gain K_p and the position of its two zeros (depending on T_i and T_d). Then, the selection of the parameter N is typically limited to pre-selected values, for example $N = 5, 10, 100, \ldots$. The consideration presented in Chapter 8 remains valid, but the frequency response of the loop transfer function $L(s)$ has to be designed considering that the controller will just add two poles (one of which in the origin, the other at high frequency) and two zeros.

Example 57. *[Tuning of a PID controller] Consider a LTI SISO system characterised by the following transfer function*

$$G(s) = \frac{e^{-s}}{(1+7s)(1+21s)}$$

and assume a PID controller has to be tuned to obtain the following requirements for the closed-loop system:

1. *asymptotic stability;*

2. *null-steady state error for a step reference* $y^0(t) = y^0, t \geq 0$;
3. *maximum overshoot of the output* $y(t)$ *in response to a step reference of* 20%, *i.e.* $S_{\%} < 20\%$;
4. *minimum possible* 1%*-settling time.*

The second requirement is forcing the controller to have the integral part. We can then try to select a PID controller in the form

$$R(s) = K_p \frac{(1 + s\tau_1)(1 + s\tau_2)}{s\left(1 + \frac{T_d}{N}s\right)}$$

Imposing $\tau_1 = 7, \tau_2 = 21$ *(and therefore* $T_i = 28$, $T_d = 147/28 = 5.25$, *the corresponding loop transfer function results as follows*

$$L(s) = K_p \frac{e^{-s}}{s\left(1 + \frac{5.25}{N}s\right)}$$

Assuming the real pole to be way after the crossover frequency ω_c, *i.e.* $N/T_d \gg \omega_c$, *the crossover frequency can be evaluated as* $\omega_c = K_p$ *and has to be maximised to minimise the* 1%*-settling time.*

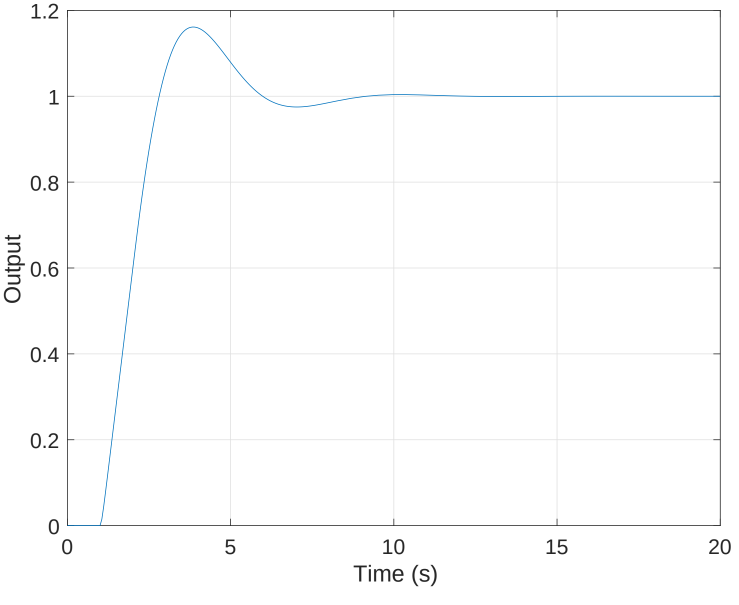


Figure 9.6: Response of the output $y(t)$ to a step reference $y^0(t)$ for the proposed tuning of the PID controller.

A lower bound for the phase margin can be evaluated from the third requirement, in fact letting

$$\xi = \sin\left(\frac{\psi_m}{2}\right)$$

the overshoot can be evaluated as

$$S_{\%} = 100e^{-\dfrac{\pi\xi}{\sqrt{1-\xi^2}}}$$

and therefore

$$S_{\%} \leq 20\% \Rightarrow \psi_m \geq 54.2521 \text{ deg}$$

In turn, since $L(s) \approx K_p e^{-s}/s$*, the phase margin can be simply evaluated as*

$$\psi_m = \pi - \left|-\frac{\pi}{2} - \omega_c\right| = \frac{\pi}{2} - \omega_c$$

Hence

$$K_p = \omega_c = \frac{\pi}{2} - \psi_m = \frac{\pi}{2} - 54.2521\frac{\pi}{180} = 0.6239 \; rad/s$$

yielding a 1%*-settling time of*

$$T_{s1\%} \approx \frac{\log_e 100}{\xi\omega_c} = 16.1883 \; s$$

In order to meet the assumption of $N \gg \omega_c T_d = 2.7584$*, we can select* $N = 10$*. Correspondingly, the transfer function of the PID controller finally assumes the following form*

$$R(s) = 0.6239\frac{1 + 19s + 84s^2}{s(1 + 0.0.0442s)}$$

As a verification, the step response of the output $y(t)$ *is reported in Fig. 9.6.*

9.3 Cascade control schemes

Cascade control is a control strategy that combines two (or more) nested feedback loops, where the output of outermost controller provides the reference to the inner control loop, as shown in Fig. 9.7. It is a commonly adopted strategies in all field of automatic control, ranging from process control to mechatronic systems.

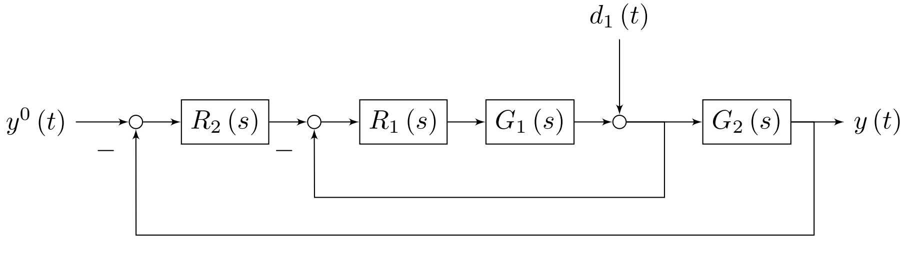


Figure 9.7: A cascade control scheme.

This strategy can be applied to systems characterised by transfer functions $G(s)$ that can be factorised in

$$G(s) = G_1(s)\,G_2(s)$$

so that $G_1(s)$ collects all high-frequency modes of the system and its output can be measured through a proper sensor. A typical mechatronic example is represented by the mass with friction of Fig. 6.1. We have already seen in Chapter 6 that the transfer function from the input force $f(t)$ to the velocity $v(t)$ can be written as

$$G_1(s) = \frac{1}{ms + r}$$

In order to also control the position of the mass, one can further consider the following transfer function

$$G_2(s) = \frac{1}{s}$$

that relates the velocity $v(t)$ and the position, say $p(t)$, of the mass. The overall transfer function, from the input force $f(t)$ to the position $p(t)$ results

$$G(s) = G_1(s) G_2(s) = \frac{1}{s(ms + r)}$$

In the following, we will investigate whether, for the general case of Fig. 9.7, the cascade control strategy can somehow outperform the alternative strategy of designing a single controller $R(s)$ for the overall system $G(s)$.

The design process starts from the inner control loop by considering transfer function $G_1(s)$ and all possible requirements to design $R_1(s)$. Assume that, as an outcome of the design process, loop transfer function $L_1(s) = R_1(s) G_1(s)$ has been obtained. Then, the design of the outer loop can start focusing on transfer function

$$F_1(s) G_2(s) = \frac{L_1(s)}{1 + L_1(s)} G_2(s)$$

which constitutes a reference for the design for the outer controller $R_2(s)$.
If the crossover frequency of the inner loop ω_{c1} has been designed sufficiently high, then

$$F_1(s) G_2(s) \approx G_2(s)$$

meaning that the design of the outer controller $R_2(s)$ can be based on transfer function $G_2(s)$, hence **neglecting the presence of the inner loop** which is a valid assumption as long as $\omega_{c2} \ll \omega_{c1}$.

As the design procedure has been sketched, we need to understand whether the cascade strategy can outperform the one with a single controller $R(s)$ for the entire system. The answer can be derived from the equivalent scheme of Fig. 9.8. The presence of the inner control loop introduces at least two important benefits:

- as already discussed, the outer controller $R_2(s)$ can be designed based on the dynamics of $G_2(s)$, only, rather than on $G(s)$;

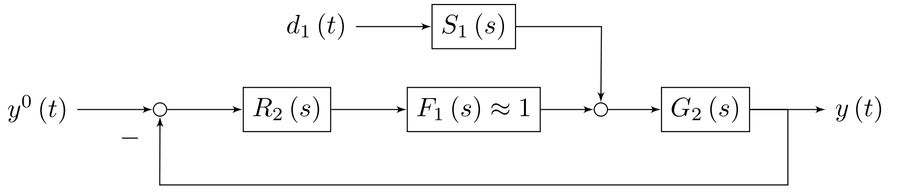


Figure 9.8: An equivalent block diagram of the cascade control scheme of Fig. 9.7.

- the presence of the inner control loop contributes in attenuating the effects of $d_1(t)$ by means of the corresponding sensitivity transfer function $S_1(s)$.

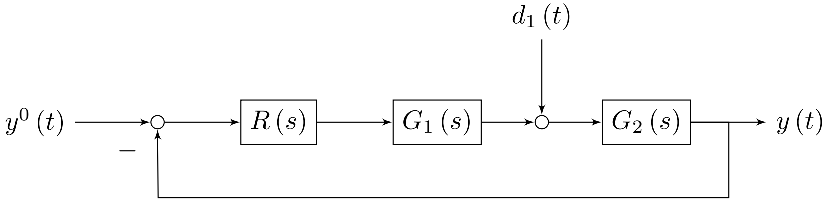


Figure 9.9: Block diagram for the a single control loop.

The situation corresponding to the cascade control loop of Fig. 9.8 has to be compared to the one of Fig. 9.9, corresponding to a single control loop with $R(s)$ designed to govern the overall system $G(s) = G_1(s)\,G_2(s)$. While in presence of the inner control loop the relationship between the disturbance $d_1(t)$ and the output $y(t)$ can be identified in the transfer function

$$\frac{S_1(s)\,G_2(s)}{1 + R_2(s)\,F_1(s)\,G_2(s)}$$

in the other case one would simply have

$$\frac{G_2(s)}{1 + R(s)\,G(s)}$$

Example 58. *[Benefits of cascaded control] With reference to the block diagram of Fig. 9.7 consider*

$$G_1(s) = \frac{1}{1 + 0.1s}, G_2(s) = \frac{1}{s}$$

and assume the following controller $R_1(s)$ has been designed based on $G_1(s)$

$$R_1(s) = 10\frac{1 + 0.1s}{s}$$

to obtain a crossover frequency $\omega_{c1} = 10$ *rad/s and phase margin* $\psi_{m1} = 90$ deg. *The outer controller* $R_2(s)$ *can be designed focusing on*

$$F_1(s)\, G_2(s) = \frac{1}{(1 + 0.1s)\, s}$$

The magnitude of its frequency response can be approximated with the one of $G_2(s) = 1/s$ *at least for frequencies* $\omega \ll \omega_{c1}$. *According to this approximation one can design*

$$R_2(s) = 1$$

which provides a crossover frequency of $\omega_{c2} = 1$ *rad/s and a phase margin of* $\psi_{m2} = 90$ deg.

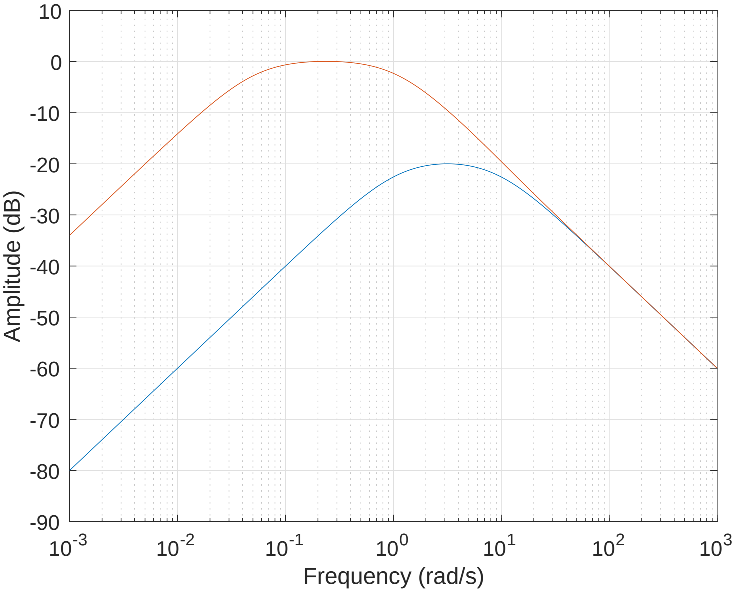


Figure 9.10: Magnitude diagrams of the frequency responses of the transfer function between the disturbance $d_1(t)$ and the output $y(t)$ for the cascade control loop (blue) and for the case of a single controller (red).

Considering the alternative strategy in Fig. 9.9, designing a single controller $R(s)$ *for the entire system* $G(s) = G_1(s)\, G_2(s)$ *to obtain roughly the same performance, one would have obtained*

$$R(s) = 0.05\frac{1 + 20s}{s}$$

which results in a crossover frequency of $\omega_c = 0.996$ *rad/s and a phase margin of* $\psi_m = 81.4$ deg. *The corresponding relationships between the reference signal* $y^0(t)$ *and the output signal* $y(t)$ *are*

$$\frac{R_2(s)\, F_1(s)\, G_2(s)}{1 + R_2(s)\, F_1(s)\, G_2(s)} \approx \frac{R_2(s)\, G_2(s)}{1 + R_2(s)\, G_2(s)} = \frac{1}{1+s}$$

and

$$\frac{R(s)\, G(s)}{1 + R(s)\, G(s)} \approx \frac{1}{1 + s/0.996}$$

respectively. Therefore, the two strategies will behave the same in terms of their output responses the same reference signal. As discussed before, it is in the response of the output $y(t)$ to disturbance $d_1(t)$ that one would expect the major difference between the two strategies. In fact, for the cascaded strategy one would have

$$\frac{S_1(s)\,G_2(s)}{1+R_2(s)\,F_1(s)\,G_2(s)} \approx \frac{S_1(s)\,G_2(s)}{1+R_2(s)\,G_2(s)} = \frac{s^2}{0.1s^2+0.1s+1}$$

while, in turn, with the single control loop

$$\frac{G_2(s)}{1+R(s)\,G(s)} \approx \frac{20s}{(1+20s)\,(1+s/0.996)}$$

Figure 9.10 reports the magnitude diagrams of the frequency responses of the transfer function between the disturbance $d_1(t)$ and the output $y(t)$ for the cascade control loop and for the case of a single controller. As one can notice, the attenuation of the effects of the disturbance $d_1(t)$ on the output $y(t)$ is much larger in the case of the cascade control structure.

As further confirmation, Fig. 9.11 finally reports the response of the output $y(t)$ to a step reference $y^0(t)$ and to a step disturbance $d_1(t)=0.1, t\geq 0$ in the two scenarios. As one can appreciate, despite the same crossover frequencies $\omega_{c2}=\omega_c$ (which can be verified from the same rising times), the cascade control loop is capable of better handling the rejection of the effects of the disturbance $d_1(t)$ on the output $y(t)$.

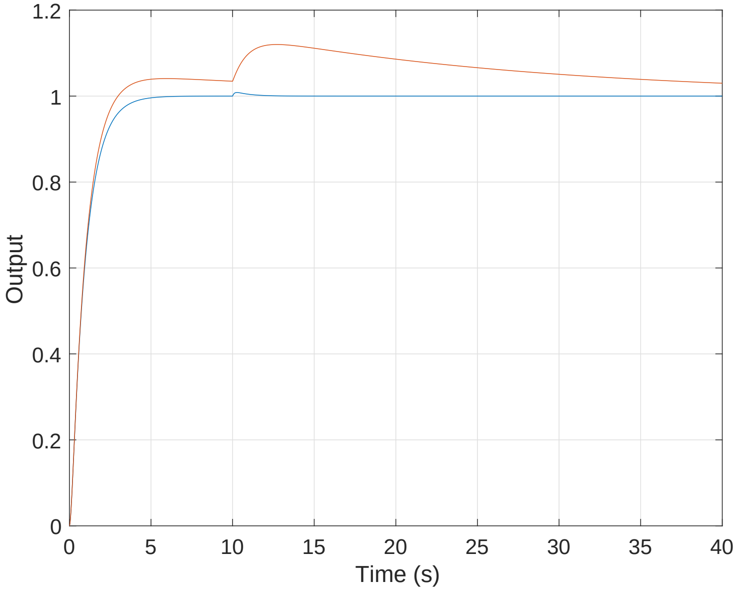


Figure 9.11: Output $y(t)$ in response to a step reference $y^0(t)$ and to a step disturbance $d_1(t)=0.1, t\geq 10$ for the cascade control loop (blue) and for the case of a single controller (red).

Example 59. *[Cascaded control for an under-actuated system] Consider an Unmanned Aerial Vehicle (UAV) in the vertical plane, as the one reported in Fig. 9.12, where $m = 2$ kg and $J = 0.0267$ $kg \cdot m^2$ are the mass and the moment of inertia, respectively, while $f_i(t)$ are the thrust forces. Moreover, let $x(t)$ and $\theta(t)$ be the horizontal displacement of the centre of mass, and the roll angle, respectively. Then, the equations governing the motion of the UAV*

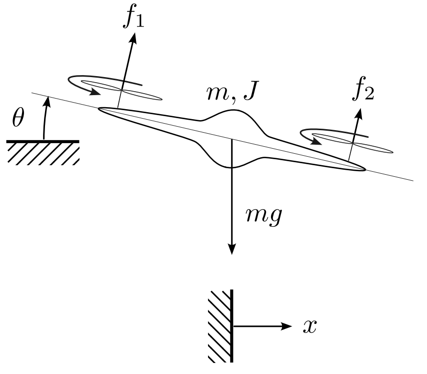


Figure 9.12: Planar model of an unmanned aerial vehicle.

can be written as follows

$$m\frac{d^2}{dt^2}z(t) = (f_1(t) + f_2(t))\cos(\theta(t)) - mg$$
$$m\frac{d^2}{dt^2}x(t) = (f_1(t) + f_2(t))\sin(\theta(t))$$
$$J\frac{d^2}{dt^2}\theta(t) = b(f_1(t) - f_2(t))$$

where $b = 0.2$ m is the distance of the propellers from the centre of gravity. For constant elevations, i.e.

$$\frac{d}{dt}z(t) = 0$$

and therefore

$$(f_1(t) + f_2(t))\cos(\theta(t)) = mg$$

and for small values of the roll angle, i.e. $\theta(t) \approx 0$, the previous equations simplify in the following ones

$$f_1(t) + f_2(t) = mg$$
$$\frac{d^2}{dt^2}x(t) = g\theta(t)$$
$$J\frac{d^2}{dt^2}\theta(t) = b(f_1(t) - f_2(t))$$

Defining $F(t) = f_1(t) + f_2(t) = mg$ and $f(t) = f_1(t) - f_2(t)$ respectively as the sum and the difference of the forces generated by the two propellers, we further obtain

$$\frac{d^2}{dt^2}x(t) = g\theta(t)$$
$$J\frac{d^2}{dt^2}\theta(t) = bf(t)$$

The block diagram corresponding to these last equations is reported in Fig. 9.13. Assuming

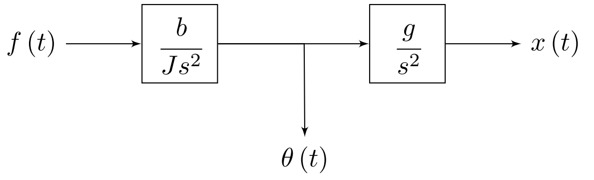


Figure 9.13: Block diagram for the control of the UAV in Fig. 9.12.

the roll angle $\theta(t)$ to be measurable, and letting

$$G_\theta(s) = \frac{b}{Js^2}, G_x(s) = \frac{g}{s^2}$$

being the transfer functions of the roll and the lateral dynamics, respectively, we can surely setup a cascade control strategy with an inner controller $R_\theta(s)$ to govern the roll dynamics and an outer one, say $R_x(s)$, to govern the lateral motion.

The inner controller $R_\theta(t)$ to govern the roll dynamics can be designed focusing on $G_\theta(s)$. As an example, using the following PID controller

$$R_\theta(s) = 0.4\frac{(1+3s)^2}{s(1+s/100)}$$

the corresponding loop transfer function $L_\theta(s) = R_\theta(s)G_\theta(s)$ corresponds to a crossover frequency $\omega_{c\theta} = 27$ rad/s and a phase margin of $\psi_m = 87$ deg. The corresponding complementary sensitivity $F_\theta(s)$ results as follows

$$F_\theta(s) = \frac{800s^2 + 1600s + 800}{0.02667s^4 + 26.67s^3 + 800s^2 + 1600s + 800}$$

and can be safely assumed to have unitary magnitude for $\omega \ll \omega_{c\theta} = 27$ rad/s. From the arguments above, we can design controller $R_x(s)$ only based on $G_x(s)$. For example, if we select the following PID controller

$$R_x(s) = 3\cdot 10^{-3}\frac{(1+10s)^2}{s(1+s/10)}$$

the corresponding (approximated) loop transfer function $L_x(s) \approx R_x(s)G_x(s)$ will have a crossover-frequency of $\omega_{cx} \approx 1$ rad/s and a phase margin of $\psi_m = 71.3$ deg. Corresponding step response for the closed-loop system are reported in Fig. 9.14.

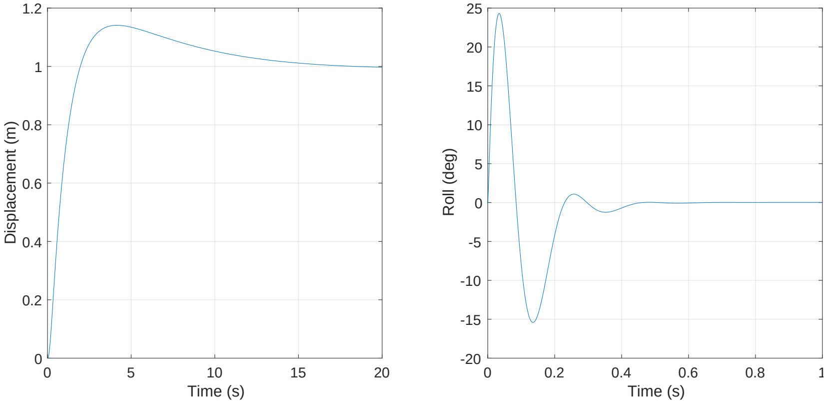


Figure 9.14: Step responses for a unit displacement $x^0(t) = 1, t \geq 0$ of the position $x(t)$ (left) and the roll angle $\theta(t)$ (right).

Chapter 10

Root locus

In the previous Chapters we have addressed the problem of designing the transfer function of the controller $R(s)$ relying on the Bode's criterion for the stability of the corresponding closed-loop system. As we have seen, the Bode's criterion is only applicable to systems $G(s)$ and controllers $R(s)$ with no poles within the open right half-plane. Hence, the synthesis method derived in Chapter 8 cannot be applied if these conditions are not satisfied.

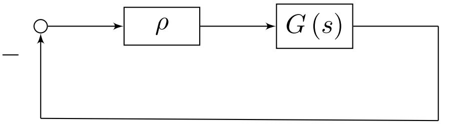


Figure 10.1: A negative feedback control system.

10.1 The root locus

Consider the negative feedback control system in Fig. 10.1 where ρ is a design parameter to be tuned and $G(s)$ stands for the transfer function of the strictly-proper system to be controlled. The loop transfer function for such a system can be simply identified as $L(s) = \rho G(s)$, hence the corresponding closed-loop characteristic equation is as follows

$$1 + L(s) = 1 + \rho G(s) = 0 \tag{10.1}$$

Definition 13. *The root locus is the graphical representation (in the complex plane) of the roots of* (10.1) *for all possible values of parameter ρ. In particular, the locus for $\rho > 0$ is call direct, the one for $\rho < 0$ is called inverse.*

Example 60. *[Second-order characteristic equation] Consider the following polynomial equation*

$$s^2 + s - 2 + \rho = 0$$

which corresponds to the closed-loop characteristic equation of the system in Fig. 10.1 when

$$G(s) = \frac{1}{(s-1)(s+2)}$$

The roots for this equation are simply computed as

$$s_{1,2} = \frac{-1 \pm \sqrt{1 - 4(\rho - 2)}}{2} = \frac{-1 \pm \sqrt{9 - 4\rho}}{2}$$

Whenever $\Delta = 9 - 4\rho \geq 0$, i.e. when $\rho \leq 9/4$ roots are real. Starting from $\rho = 0$, roots are those of the open-loop system, hence $s = 1$ and $s = -2$. By slightly increasing the value of ρ from $\rho = 0$, the roots are moving towards each other until meeting on the real axis when $\rho = 9/4$. In turn, when $\Delta < 0$, that is for $\rho > 9/4$, the roots are complex and conjugate with negative real part equal to $\mathbb{R}\text{e}[s] = -1/2$ no matter what $\rho > 9/4$ is.
When $\rho < 0$, roots are always real. One always has a negative real part, the other has always a positive real part since $-1 + \sqrt{9 - 4\rho} < 0$ when $\rho < 0$.

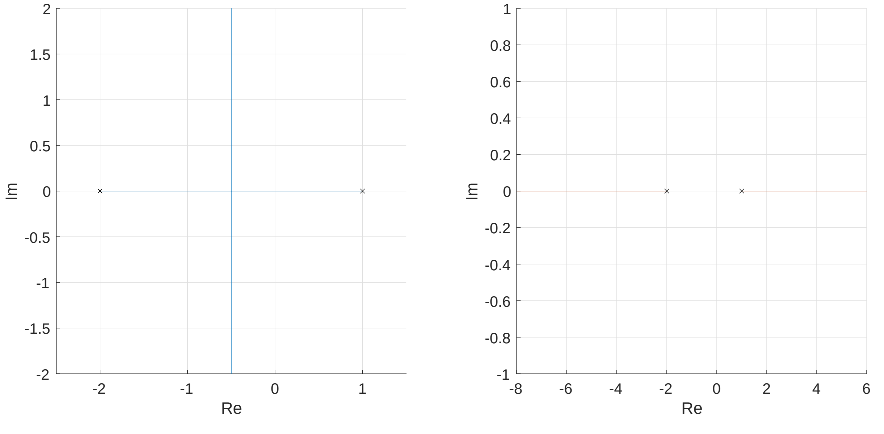


Figure 10.2: Direct (left) and inverse (right) loci for the system considered in this example.

In the previous example we have seen that plotting the direct and inverse loci is relatively simple for second-order systems. However, as we have seen in Chapter 2, the calculation of closed-loop roots, especially when they are function of a parameter (as in the case of the root locus), is usually prohibitive, if not impossible. For this reason, we are looking for alternatives to sketch the root locus. A set of drawing rules are available and listed in the following.

Property 16. *Given the closed-loop system in Fig. 10.1 with*

$$G(s) = \frac{\prod_i (s + z_i)}{\prod_j (s + p_j)}$$

1. *the locus is symmetric with respect to the real axis;*
2. *the root locus is always composed of n branches, where n is the order of transfer function $G(s)$;*
3. *the locus has ν asymptotes (being ν the relative degree);*
4. *all the asymptotes meet on the real axis at coordinate*

$$x_a = \frac{\sum_i z_i - \sum_j p_j}{\nu}$$

5. *the direction of each asymptote $h = 0, \ldots, \nu - 1$ can be computed considering its angle θ_a with respect to the positive real axis, where*

$$\theta_a = \begin{cases} \dfrac{\pi + h2\pi}{\nu}, \rho > 0 \\ \dfrac{h2\pi}{\nu}, \rho < 0 \end{cases}$$

6. *each branch starts ($\rho = 0$) from an open-loop pole of $G(s)$, i.e. from $s = -p_j$ and terminates ($\rho \to \pm\infty$) following one of the asymptotes;*
7. *all points on the real axis belong either to the direct or to the inverse locus, in particular those leaving on their right an odd number of singularities (poles or zeros) of $G(s)$ belong to the direct locus.*

Furthermore, if $s \in \mathbb{C}$ is a point on the locus (either direct or inverse), then it necessarily solves

$$1 + \rho G(s) = 0$$

and therefore

$$|\rho| = \frac{1}{|G(s)|}$$

Example 61. *[Example of root locus] Consider*

$$G(s) = \frac{s - 1}{(s + 1)(s + 2)}$$

the corresponding root locus has $n = 2$ branches and $\nu = 1$ asymptote with

$$\theta_a = \begin{cases} \pi, \rho > 0 \\ 0, \rho < 0 \end{cases}$$

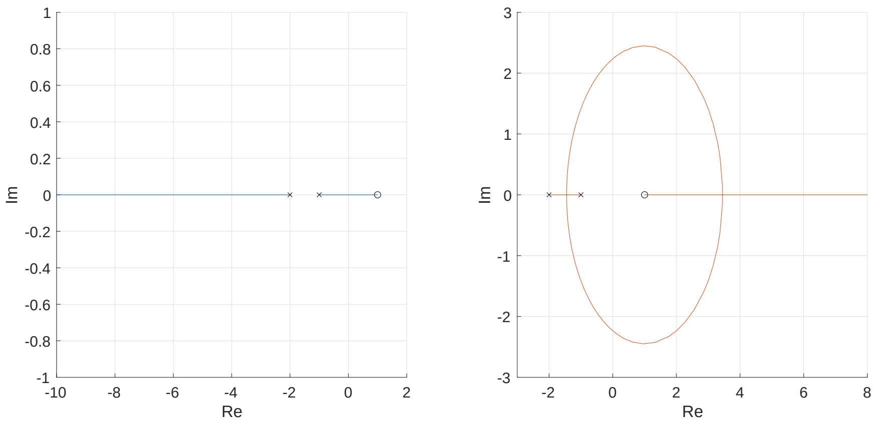


Figure 10.3: Direct (left) and inverse (right) loci for the system considered in this example.

The drawing of the inverse locus (see Fig. 10.3(right) deserves particular attention. Since all points s on the real axis such that $-2 < s < -1$ have to belong to the inverse locus while $s = -2$ and $s = -1$ are the two open-loop roots, a simple line connecting the two cannot justify the presence of two branches. Therefore, the two branches generated by the two open-loop roots have to breakaway from the real axis, move towards and finally enter the right half-plane. Then the two branches meet again on the positive real axis and each branch will proceed either towards the right half-plane zero in $s = 1$ or towards the asymptote having $\theta_a = 0$.

Finally, looking at the direct locus only one can notice that one of the two branches crosses the imaginary axis in the origin, i.e. for $s = 0$. Correspondingly

$$|\rho| = \frac{1}{|G(0)|} = \frac{1}{2}$$

Since point $s = 0$ belongs to the direct locus we can safely consider the closed-loop system to be asymptotically stable at least for $0 \leq \rho < 1/2$.

Another important aspects of the root locus, that is valid only whenever the relative degree $\nu \geq 2$ is the following one.

Property 17 (Conservation of the sum)**.** *Assume that the relative degree of $G(s)$ is at least 2, i.e. $\nu \geq 2$, then the sum Σ of all the roots of $1 + \rho G(s) = 0$ is constant, i.e.*

$$\frac{d\Sigma}{d\rho} = 0, \forall \rho$$

From this property, we can observe that whenever a root of the closed-loop system moves towards the right, i.e. toward the right half-plane, e.g. for increasing values of ρ, another one has to move left. Notice that the sum of roots can be actually evaluated as the sum of their real parts, since closed-loop roots are always complex and conjugate, when not real.

Example 62. *Consider transfer function*

$$G(s) = \frac{1}{(s+1)(s+2)(s+3)}$$

then since its relative degree is $\nu = 3 \geq 2$, the aforementioned property holds and, no matter what ρ, is the sum of the roots of

$$1 + \rho G(s) = 0$$

is constant. The direct locus of $G(s)$ is reported in Fig. 10.4. As one can notice, for $\rho = 58.5$,

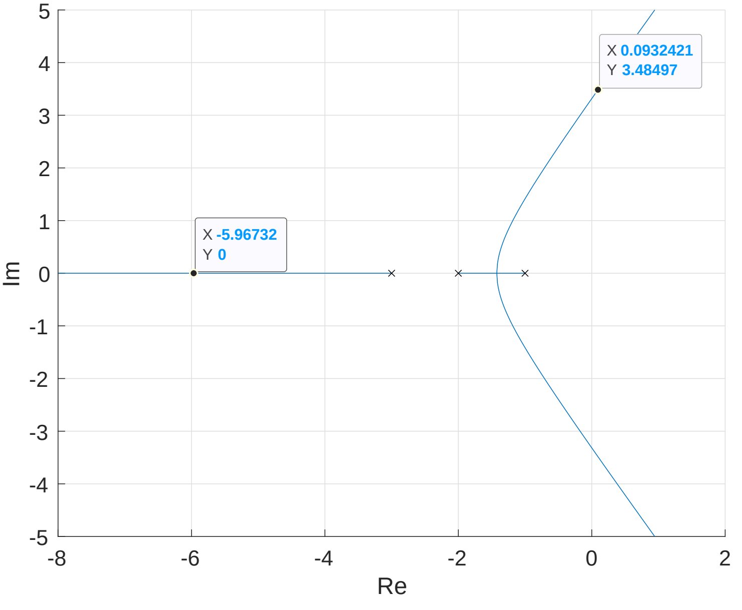


Figure 10.4: Direct locus for the system considered in this example.

the leftmost root is equal to $s = -5.97$, while the other two are in $s = -0.0163 \pm 3.29j$. Their sum Σ can be simply evaluated as

$$\Sigma = -5.97 - 0.0163 - 0.0163 = -6$$

which is the sum of open-loop roots (i.e. for $\rho = 0$).

10.2 Synthesis with the root locus method

Differently from the design method relying on the magnitude diagram of the frequency response of $L(s)$ addressed in Chapter 8, for the root locus method there are no specific guidelines for the synthesis of the controller $R(s)$ which has to be selected in the form

$$R(s) = \frac{\prod_i (s + z_i)}{\prod_j (s + p_j)}$$

Figure 10.5 reports the block diagram of the closed-loop system as a reference for the synthesis of the controller which comprises both transfer function $R(s)$ and parameter ρ. Clearly, for given $G(s)$ one has to design $R(s)$ and then

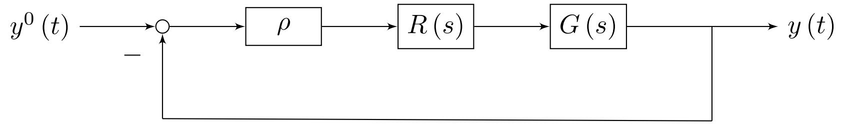


Figure 10.5: A negative feedback control system for the synthesis.

tune parameter ρ applying the root locus method to the closed-loop characteristic equation

$$1 + \rho L(s) = 0$$

where obviously $L(s) = R(s)G(s)$. As previously mentioned, there are no specific guidelines on how to design transfer function $R(s)$. On the other hand, poles and zeros in $R(s)$, and therefore in $L(s)$, will influence the corresponding root locus.

Example 63. *[Synthesis with the root locus method] Consider the following system to be controlled*

$$G(s) = \frac{1}{s-1}$$

and assume the following requirements for closed-loop system in response to the reference signal $y^0(t) = 1, \forall t > 0$:

1. *the steady-state error is null;*
2. *the settling time is not higher than 2 seconds;*
3. *the maximum overshoot is of 5%.*

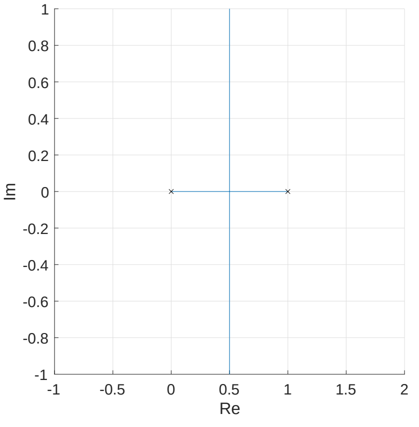
To handle the first requirement on the steady-state error, we should have learned from the findings of Chapter 8 that a pole in the origin is needed in the transfer function of the controller, hence we will consider

$$R(s) = \frac{1}{s} \frac{\prod_i (s + z_i)}{\prod_j (s + p_j)}$$

where z_i's and p_j's have to be determined. Figure 10.6 reports the direct and inverse loci for the preliminary design of the controller corresponding to $R(s) = 1/s$. As one can see, as both the loci have at least one of the two branches constantly in the right half-plane, it is not possible to make the closed-loop system asymptotically stable, neither with $\rho > 0$ nor with $\rho < 0$.

As per the other two requirements, we can try to design the controller so to achieve the following closed-loop characteristic polynomial

$$s^2 + \xi\omega_n s + \omega_n^2$$



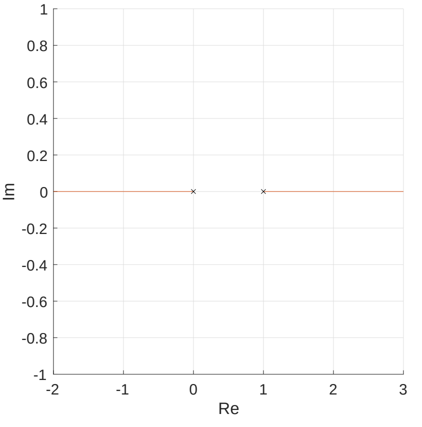


Figure 10.6: Direct (left) and inverse (right) loci for the preliminary controller $R(s) = 1/s$.

where ξ *can be obtained to enforce the maximum overshoot* $S_{\%}$ *to be lower than the prescribed value* ($S_{\%} \leq 5\%$), *i.e.*

$$\xi \geq 0.6901$$

As the damping ξ *is the cosine of the angle* α *formed by the complex root with the negative real axis, we can alternatively require*

$$\alpha \leq \arccos(\xi) = 46.3614\,\mathrm{deg}$$

Moreover, in order to obtain a settling time lower than the one required, i.e. $T_{s,1\%} \leq 2$, *we need*

$$\frac{4.6}{\xi\omega_n} \leq 2$$

Since $-\xi\omega_n$ *is the real part of the complex number having magnitude* ω_n *and damping* ξ, *the requirement on the settling time can be easily translated into*

$$\mathbb{R}e\,[s] = -\xi\omega_n \leq -\frac{4.6}{2}$$

By combining the two requirements, we can easily identify the region in the complex plane where closed-loop roots are expected to be placed (see Fig. 10.7).

In order to have the root locus of $L(s)$ *to enter somehow in the highlighted region in Fig. 10.7, we can decide to add a left half-plane zero in the controller. This way, we are changing the relative degree of the loop transfer function to* $\nu = 1$. *This way, only one asymptote will be expected corresponding to*

$$\theta_a = \frac{\pi}{v} = \pi$$

for the direct locus $\rho > 0$. *As the loop transfer function will be still a second order one, the two branches will tend either to the asymptote (in the left half-plane) or to the left half-plane zero. Let us consider, for example*

$$R(s) = \frac{s+2}{s}$$

The corresponding direct root locus is shown in Fig. 10.8. Eventually, for some value of $\rho > \rho_{\min}$, *both the branches are inside the highlighted region, and specifically* $\rho_{\min} \approx 5.6$. *Notice that the overall transfer function*

$$\rho R(s) = \rho\frac{s+2}{s} = \frac{\rho s + 2\rho}{s} = \rho + \frac{2\rho}{s}$$

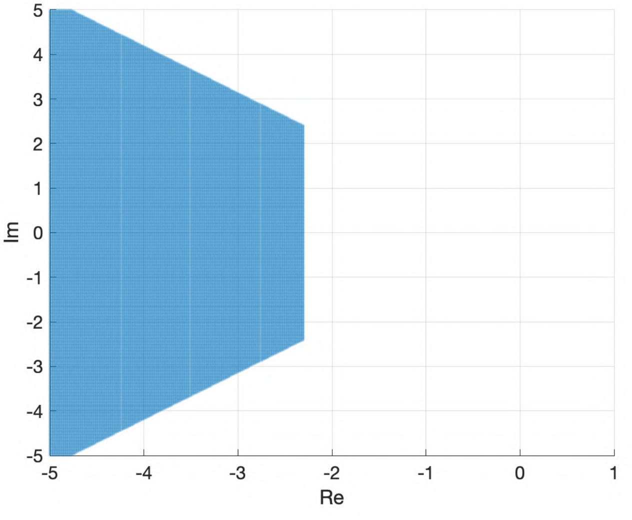


Figure 10.7: Region in the complex plane hosting closed-loop roots compatible with the given requirements.

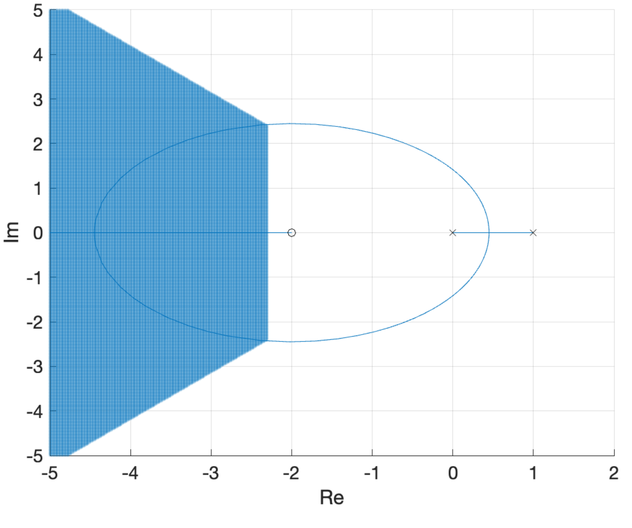


Figure 10.8: Direct root locus corresponding to the controller $R(s) = (s+2)/s$.

actually corresponds to the one of a proportional-integral controller with $K_p = \rho$ and $K_i = 2\rho$. The step response of the closed-loop system, for $\rho = 6$ and considering a feed-forward compensator $C(s) = -6$ is reported in Fig. 10.9.

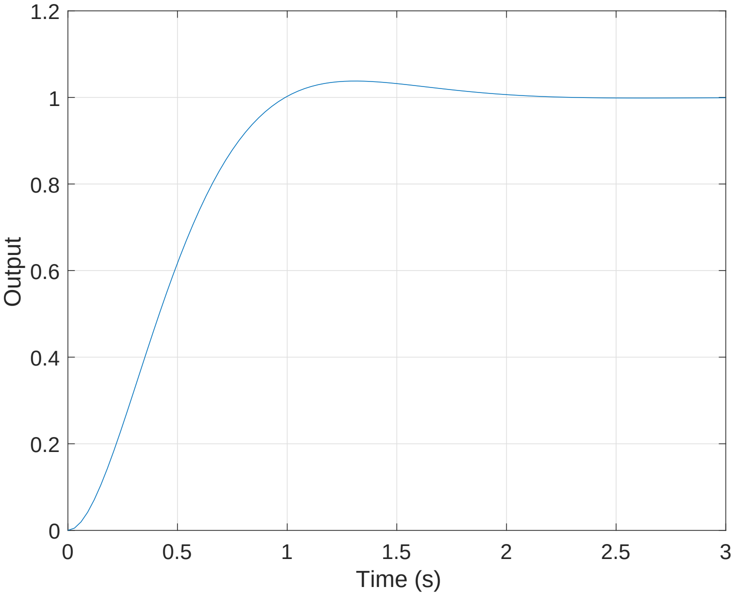


Figure 10.9: Step response of the closed-loop system for a proportional-integral controller $R(s) = 6 + 12/s$ and a feed-forward compensator $C(s) = -6$.

10.3 Contour locus method

As of now, this Chapter discussed the root locus method to derive the position of the closed-loop roots of the system in Fig. 10.1 by varying parameter ρ. In the last example, we have seen this method applied to support the design of the controller $R(s)$. In both the cases, the direct and inverse loci graphically represented the positions in the complex plane of the solutions of the characteristic equation

$$1 + \rho L(s) = 0$$

More generally, assume that we are now interested in analysing the position of the closed-loop roots of the system in Fig. 10.10 which are the solutions of the characteristic following equation

$$1 + L(\rho, s) = 0$$

In order to still rely on the root locus method, assume that the closed-loop characteristic equation can be written as

$$0 = \rho a(s) + b(s)$$

where $a(s)$ and $b(s)$ are polynomial functions of $s \in \mathbb{C}$. Dividing both members

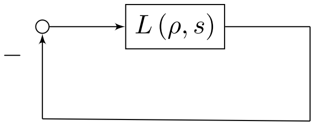


Figure 10.10: A generic negative feedback control system.

by $b(s)$, one obtains

$$0 = 1 + \rho \frac{a(s)}{b(s)}$$

Provided that $a(s)/b(s)$ is a strictly proper transfer function, the root locus method can be still applied. Otherwise, one can define $k = 1/\rho$ and apply the root locus method to

$$0 = 1 + k \frac{b(s)}{a(s)}$$

Example 64. *[Contour locus for the selection of the integral gain] Consider the following transfer function*

$$G(s) = \frac{1}{1+s}$$

which represents the system to be controller by the following proportional-integral controller

$$R(s) = K_p + \frac{K_i}{s}, K_p > 0, K_i > 0$$

For values of K_i such that the zero of the controller at frequency $1/T_i$ anticipates the crossover frequency ω_c (i.e. $1/T_i < \omega_c$), the loop transfer function can be approximated as

$$R(s)\,G(s) = \frac{K_p s + K_i}{s}\frac{1}{1+s} \approx \frac{K_p}{s}$$

and therefore the proportional gain K_p can be selected to properly tune the crossover frequency $\omega_c = K_p$, hence $1/T_i = K_i/K_p < K_p$ and therefore $K_i < K_p^2$. Assume, for example $K_p = 100$ (therefore $K_i < 10000$ will be assumed). Correspondingly, the corresponding closed-loop characteristic equation

$$0 = 1 + L(K_i, s) = 1 + \frac{100s + K_i}{s(s+1)}$$

depends on the integral gain K_i which is still to be tuned. One can easily rewrite the closed-loop characteristic equation as follows

$$0 = s(s+1) + 100s + K_i = s^2 + 101s + K_i = 1 + K_i \frac{1}{s^2 + 101s} = 1 + K_i \frac{1}{s(s+101)}$$

which is now in the form for the applicability of the root locus method with respect to the integral gain K_i.

From Fig. 10.11 one can notice that closed-loop roots are initially (for small values of K_i) real and then become complex and conjugate. In particular, the vertical asymptote is characterised by

$$x_a = \frac{0 - (0 + 101)}{1} = -50.5$$

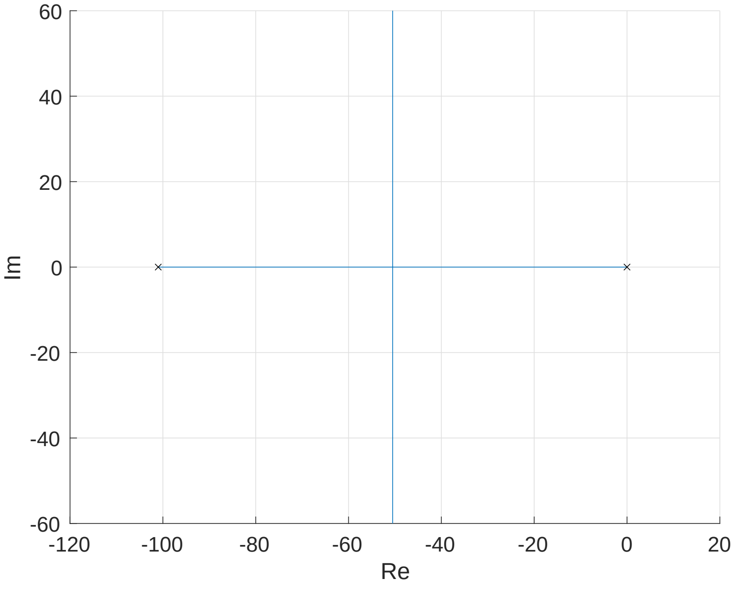


Figure 10.11: Direct root locus for varying integral gain $K_i > 0$.

which corresponds to $K_i = |-50.5 + 0|\,|-50.5 + 101| = 2550.3$. *This means that for* $K_i = 2550.3$, *the closed-loop system will be characterised by the following characteristic equation*

$$0 = (s + 50.5)^2$$

The reported plots should not be surprising. Referring to the root locus in Fig. 10.11, for $K_i = 100$ *one of the two closed-loop poles is still closed to the open-loop position* $s = 0$ *(actually* $s = -1$*) which is justifying the slow behaviour of the output in response to the disturbance* $w(t)$.

From a different perspective, for $K_i = 100$, *the zero of the controller has been placed to cancel out the pole of* $G(s)$ *in* $s = -1$ *which happens two decades before the crossover-frequency* $\omega_c = 100$ *rad/s. While corresponding to the tuning obtained using the root locus method, i.e. for* $K_i = 2550.3$, *the zero of the proportional-integral controller has been set to frequency* $\omega_z \approx 25.5$ *rad/s* $\approx \omega_c/4$, *hence slightly before the crossover-frequency. Figure 10.13 reports the Bode's diagrams of the frequency response of the loop transfer function* $L(s)$ *in the two cases.*

As one can see, both the tunings of the proportional-integral controller attains a crossover-frequency of $\omega_c \approx 100$ *rad/s and a phase margin near* 90 deg. *On the other hand, the tuning performed with the root locus method is characterised by a higher magnitude of* $L(j\omega)$ *for* $\omega \ll \omega_c$ *which justifies the better attenuation of the disturbance* $w(t)$.

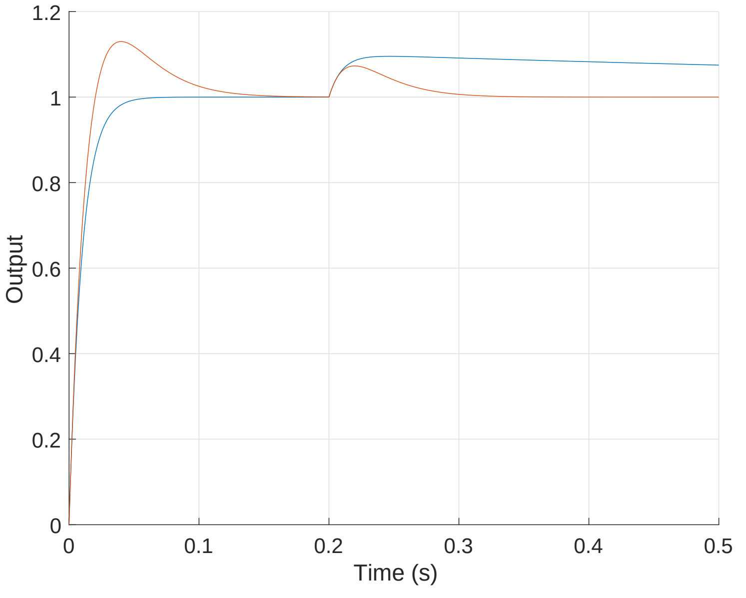


Figure 10.12: Responses of the closed-loop system to a step reference $y^0(t)$ and a step disturbance $w(t) = 10, t \geq 0.2$ with the proportional-integral controllers $R(s) = 100 + K_i/s$, for $K_i = 100$ (blue), and $K_i = 2550.3$ (red).

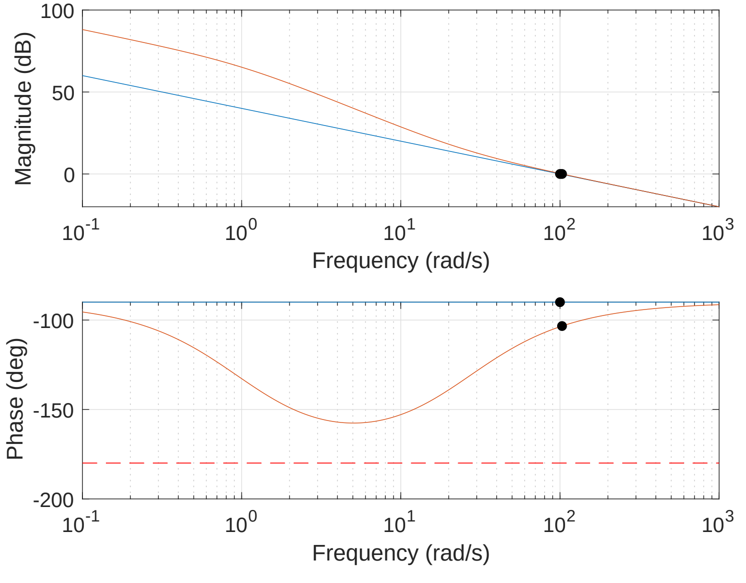


Figure 10.13: Bode's diagrams of the frequency response of the loop transfer function for $K_i = 100$ (blue), and $K_i = 2550.3$ (red).

Chapter 11

Control systems design in state-space

In the previous Chapter, we have introduced the root locus method to understand how parameter ρ of the loop transfer function influences the position of closed-loop poles. We have seen that, with a proper tuning of ρ, closed-loop roots could be placed in points along the locus. The loci could be modified by inserting in the loop transfer function $R(s)$, so to open for new placement possibilities. Unfortunately, no guidelines could be inferred on how to design such a transfer function $R(s)$. This Chapter is intended to fill this gap by proposing a method to design the transfer function of the controller $R(s)$ based on specified position of closed-loop poles.

11.1 State feedback controller design

Consider a LTI single-input system

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}u(t)$$

and assume its state to be completely accessible and measurable as an output of the system, i.e.

$$\boldsymbol{y}(t) = \boldsymbol{x}(t)$$

Then, using following state feedback control law

$$u(t) = \boldsymbol{K}\boldsymbol{y}(t) = \boldsymbol{K}\boldsymbol{x}(t) \tag{11.1}$$

By substituting the control law in (11.1) in the dynamics of the system, one obtains the following closed-loop system

$$\frac{d}{dt}\boldsymbol{x}(t) = (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{K})\,\boldsymbol{x}(t)$$

Assume we want to select the control gain $\boldsymbol{K}$ so that the closed-loop system has is characterised by the following characteristic polynomial

$$\prod_{i=1}^{n}(s+\lambda_i)$$

where λ_i's represent some desired position of the closed-loop eigenvalues. Depending on $\boldsymbol{K}$, the closed-loop eigenvalues are the roots of the characteristic polynomial

$$\det\left(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{B}\boldsymbol{K}\right)$$

If one is interested in designing $\boldsymbol{K}$ so to place the eigenvalues of the closed-loop system corresponding to the specified λ_i's, one has to actually make the two characteristic polynomial identical, i.e.

$$\det\left(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{B}\boldsymbol{K}\right)=\prod_{i=1}^{n}(s+\lambda_i)$$

which can be done by constraining their coefficients to be the same.

Example 65. *[Mass-spring-damper system] Consider a generic mass-spring-damper system as the one in Fig. 2.2. For $f(t)=k_1x_1(t)+k_2x_2(t)$, the closed-loop characteristic polynomial results as follows*

$$s^2+\frac{r-k_2}{m}s+\frac{k-k_1}{m}$$

in fact

$$\begin{aligned}\frac{d}{dt}x_1(t)&=x_2(t)\\ \frac{d}{dt}x_2(t)&=\frac{k_1-k}{m}x_1(t)+\frac{k_2-r}{m}x_2(t)\end{aligned}\tag{11.2}$$

Then, assuming one wants to have closed-loop eigenvalues in $s=-7$ and $s=-8$, that is

$$(s+8)(s+7)=s^2+15s+56$$

we obtain the two following equations

$$\frac{r-k_2}{m}=15,\ \frac{k-k_1}{m}=56$$

to be solved for k_1 and k_2. Notice that, since

$$x_2(t)=\frac{d}{dt}x_1(t)$$

the control action can be rewritten as

$$u(t)=k_1x_1(t)+k_2\frac{d}{dt}x_1(t)$$

which represents the control law of an ideal ISA proportional-derivative controller with

$$K_p=-k_1, K_d=-k_2, b=0, c=0$$

We now wonder whether the system of equations to be solved to impose

$$\det\left(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{B}\boldsymbol{K}\right)=\Pi_{i=1}^{n}(s+\lambda_i)$$

has exactly one solution (or at least one).

Example 66. *[RLL circuit] Consider the RLL circuit of Fig. 2.9 where*

$$\boldsymbol{A}=-\frac{R}{L}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix},\boldsymbol{B}=\frac{R}{L}\begin{bmatrix}1\\ 1\end{bmatrix}$$

and assume the input current to be set as $i\left(t\right)=k_1x_1\left(t\right)+k_2x_2\left(t\right)$ The corresponding closed-loop system will be characterised by matrix

$$\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}=\begin{bmatrix}\frac{R\left(k_1-1\right)}{L} & \frac{R\left(k_2-1\right)}{L}\\ \frac{R\left(k_1-1\right)}{L} & \frac{R\left(k_2-1\right)}{L}\end{bmatrix}$$

the corresponding characteristic polynomial results

$$\det\left(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{B}\boldsymbol{K}\right)=s^2+\frac{2R-k_1R-k_2R}{L}s$$

Assuming s^2+bs+c to be the desired characteristic polynomial for the closed-loop system, one has to solve for k_1 and k_2 the following system of equations

$$\begin{bmatrix}-R & -R\\ 0 & 0\end{bmatrix}\begin{bmatrix}k_1\\ k_2\end{bmatrix}=\begin{bmatrix}bL-2R\\ c\end{bmatrix}$$

for which the existence of a solution cannot be always guaranteed.

In the light of the outcome of the previous example, we are introducing the following result.

Property 18 (Pole placement)**.** *The problem of designing $\boldsymbol{K}$ based on an arbitrary specification of the position of the eigenvalues of $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ admits a solution if and only if the pair $(\boldsymbol{A},\boldsymbol{B})$ is reachable.*

11.2 State observer design

In the previous Section, we have seen that provided that the pair $(\boldsymbol{A},\boldsymbol{B})$ is reachable it is possible to arbitrarily assign the eigenvalues of $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ with a proper selection of the control parameter $\boldsymbol{K}$. Unfortunately, the method is applicable with the prohibitive requirement of the state vector $\boldsymbol{x}\left(t\right)$ being completely measurable, i.e. only when the output equation ot the LTI system to be controlled is

$$\boldsymbol{y}\left(t\right)=\boldsymbol{x}\left(t\right)$$

We know want to understand whether for a more generic strictly-proper LTI SISO system

$$\begin{aligned}\frac{d}{dt}\boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}u(t) \\ y(t) &= \boldsymbol{C}\boldsymbol{x}(t)\end{aligned} \tag{11.3}$$

we can still somehow apply the method. One possible workaround is to used the knowledge of the dynamics of the system to estimate the state vector and to use this estimate, say $\hat{\boldsymbol{x}}(t)$, to feed the controller, i.e. using

$$u(t) = \boldsymbol{K}\hat{\boldsymbol{x}}(t) \tag{11.4}$$

instead of $u(t) = \boldsymbol{K}\boldsymbol{x}(t)$ as considered previously.
As a preliminary idea, one can consider the Lagrange's formula and obtain and estimate $\hat{\boldsymbol{x}}(t)$ of the state $\boldsymbol{x}(t)$ as follows

$$\hat{\boldsymbol{x}}(t) = e^{\boldsymbol{A}t}\hat{\boldsymbol{x}}(0) + \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}u(\tau)\,d\tau$$

where $\hat{\boldsymbol{x}}(0)$ represents an estimate of the initial condition $\boldsymbol{x}(0)$ of the real system. As the actual state $\boldsymbol{x}(t)$ evolves according to the Lagrange's formula as

$$\boldsymbol{x}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}(0) + \int_0^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}u(\tau)\,d\tau$$

the estimation error $\boldsymbol{\varepsilon}(t) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)$ will evolve according to

$$\boldsymbol{\varepsilon}(t) = e^{\boldsymbol{A}t}\left(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}(0)\right)$$

meaning that $\lim_{t\to+\infty}\boldsymbol{\varepsilon}(t)$ will be null only if the system in (11.3) is asymptotically stable. Moreover, notice that the evolution of the estimation error $\boldsymbol{\varepsilon}(t)$ will highly depend on the eigenvalues of matrix $\boldsymbol{A}$. Therefore, if the system at hands is asymptotically stable with particularly slow transients, the estimation error will slowly tend to zero. In the light of the discussion above, we introduce the following slightly modified version of the estimator

$$\begin{aligned}\frac{d}{dt}\hat{\boldsymbol{x}}(t) &= \boldsymbol{A}\hat{\boldsymbol{x}}(t) + \boldsymbol{B}u(t) + \boldsymbol{L}\left(\hat{y}(t) - y(t)\right) = \\ &= \left(\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C}\right)\hat{\boldsymbol{x}}(t) + \boldsymbol{B}u(t) - \boldsymbol{L}y(t) \\ \hat{y}(t) &= \boldsymbol{C}\hat{\boldsymbol{x}}(t)\end{aligned} \tag{11.5}$$

which is called **Luenberger's state observer** and depends on the parameter $\boldsymbol{L}$ to be selected. Notice that, differently from the naive estimator of the state that we introduced previously, the Luenberger observer is fed with the difference between the estimated output $\hat{y}(t) = \boldsymbol{C}\hat{\boldsymbol{x}}(t)$ and the measured output of the system $y(t) = \boldsymbol{C}\boldsymbol{x}(t)$, and this difference is used to slightly alter the dynamics of the estimator from those of the system in (11.3).

Given the Luenberger observer in (11.5), we want to understand the behaviour of the estimation error $\boldsymbol{\varepsilon}(t)$, which results as follows

$$\begin{aligned}\frac{d}{dt}\boldsymbol{\varepsilon}(t) &= \frac{d}{dt}(\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)) = \\ &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}u(t) - (\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C})\hat{\boldsymbol{x}}(t) - \boldsymbol{B}u(t) + \boldsymbol{L}\boldsymbol{C}\boldsymbol{x}(t) = \\ &= (\boldsymbol{A} + \boldsymbol{L}\boldsymbol{C})\boldsymbol{\varepsilon}(t)\end{aligned} \tag{11.6}$$

Provided that matrix $\boldsymbol{A}+\boldsymbol{L}\boldsymbol{C}$ has all eigenvalues within the left half-plane, the dynamics of the estimation error $\boldsymbol{\varepsilon}(t)$ is asymptotically stable, and its transients can be made arbitrarily fast by acting on the design parameter $\boldsymbol{L}$. The problem is now to understand under which condition the design parameter $\boldsymbol{L}$ can be selected to position the eigenvalues of $\boldsymbol{A}+\boldsymbol{L}\boldsymbol{C}$.
Notice that the eigenvalues of $\boldsymbol{A}+\boldsymbol{L}\boldsymbol{C}$ are those of $\boldsymbol{A}^T+\boldsymbol{C}^T\boldsymbol{L}^T$. Furthermore, in the previous Section we have introduced the possibility to design $\boldsymbol{K}$ to arbitrarily place the eigenvalues of $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$, provided that the pair $(\boldsymbol{A},\boldsymbol{B})$ was reachable. As a conclusion, we are able to select the design parameter $\boldsymbol{L}$ so to arbitrarily place the eigenvalues of $\boldsymbol{A}+\boldsymbol{L}\boldsymbol{C}$, provided that the pair $\left(\boldsymbol{A}^T,\boldsymbol{C}^T\right)$ is reachable.

Property 19 (Duality)**.** *The pair $\left(\boldsymbol{A}^T,\boldsymbol{C}^T\right)$ is reachable if and only if the pair $(\boldsymbol{A},\boldsymbol{C})$ is observable.*

Proof. The reachability matrix associated to the pair $\left(\boldsymbol{A}^T,\boldsymbol{C}^T\right)$

$$\boldsymbol{K}_R = \begin{bmatrix}\boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T & \dots & \left(\boldsymbol{A}^T\right)^{n-1}\boldsymbol{C}^T\end{bmatrix}$$

is actually the observability matrix of the pair $(\boldsymbol{A},\boldsymbol{C})$. □

In the light of the property above, we can introduce the following result.

Property 20 (Design of the Luenberger observer)**.** *The problem of designing $\boldsymbol{L}$ based on an arbitrary specification of the position of the eigenvalues of $\boldsymbol{A}+\boldsymbol{L}\boldsymbol{C}$ admits a solution if and only if the pair $(\boldsymbol{A},\boldsymbol{C})$ is observable.*

Example 67. *[Mass-spring-damper system] Consider again the generic mass-spring-damper system of Fig. 2.2 and assume that only its position $x_1(t)$ can be measured.*

$$\begin{aligned}\frac{d}{dt}x_1(t) &= x_2(t)\\ \frac{d}{dt}x_2(t) &= -\frac{k}{m}x_1(t) + -\frac{r}{m}x_2(t)\\ y(t) &= x_1(t)\end{aligned} \tag{11.7}$$

Then, assuming one wants to construct a Luenberger observer to estimate its full state $\boldsymbol{x}(t)$ and that the eigenvalues of the dynamics of the reconstruction error have to be placed to be 10 faster than those of the mass-spring-damper.
Since

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

and the pair $(\boldsymbol{A}, \boldsymbol{C})$ is observable, we can start finding the parameter of the observer $\boldsymbol{L}$ which places both the eigenvalues of $\boldsymbol{A} + \boldsymbol{LC}$ in $10\omega_n = 10\sqrt{k/m}$. Correspondingly, the characteristic polynomial of matrix $\boldsymbol{A} + \boldsymbol{LC}$ is expected to be as follows

$$\det\left(s\boldsymbol{I} - \boldsymbol{A} - \boldsymbol{LC}\right) = \left(s + 10\sqrt{\frac{k}{m}}\right)^2 = s^2 + 20\sqrt{\frac{k}{m}} + 100\frac{k}{m}$$

On the other hand, given $\boldsymbol{L} = \begin{bmatrix} l_1 & l_2 \end{bmatrix}^T$ the characteristic polynomial of $\boldsymbol{A} + \boldsymbol{LC}$ can be also written as

$$\det\left(s\boldsymbol{I} - \boldsymbol{A} - \boldsymbol{LC}\right) = s^2 + \frac{r - l_1 m}{m}s + \frac{k - l_1 r - l_2 m}{m}$$

The solution can be finally found solving for l_2 and l_2 the two following equations

$$l_1 = \frac{r}{m} - 20\sqrt{\frac{k}{m}}, l_1 r + l_2 m = -99k$$

11.3 Output feedback controller design

As we have seen, provided that the pair $(\boldsymbol{A}, \boldsymbol{B})$ is reachable, it is possible to arbitrarily assign the eigenvalues of $\boldsymbol{A} + \boldsymbol{BK}$ with a proper selection of the control parameter $\boldsymbol{K}$. On the other hand, if the pair $(\boldsymbol{A}, \boldsymbol{C})$ is observable, we can design the parameter $\boldsymbol{L}$ of the Luenberger observer to arbitrarily design the position of the eigenvalues of $\boldsymbol{A} + \boldsymbol{LC}$, which govern the dynamics of the estimation error.
We now wonder whether we can safely combine a Luenberger observer to estimate the state $\boldsymbol{x}(t)$ of the system in (11.3) and to adopt the control law in (11.4) as a feedback controller. By combining the two strategies to (11.3), we obtain the following closed-loop system

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{BK}\hat{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{BK}\left(\boldsymbol{x}(t) - \boldsymbol{\varepsilon}(t)\right) \\ \frac{d}{dt}\boldsymbol{\varepsilon}(t) &= \left(\boldsymbol{A} + \boldsymbol{LC}\right)\boldsymbol{\varepsilon}(t) \end{aligned} \tag{11.8}$$

which can be rewritten, in matrix form, as follows

$$\frac{d}{dt}\begin{bmatrix} \boldsymbol{x}(t) \\ \boldsymbol{\varepsilon}(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} + \boldsymbol{BK} & -\boldsymbol{BK} \\ \boldsymbol{0} & \boldsymbol{A} + \boldsymbol{LC} \end{bmatrix}\begin{bmatrix} \boldsymbol{x}(t) \\ \boldsymbol{\varepsilon}(t) \end{bmatrix} \tag{11.9}$$

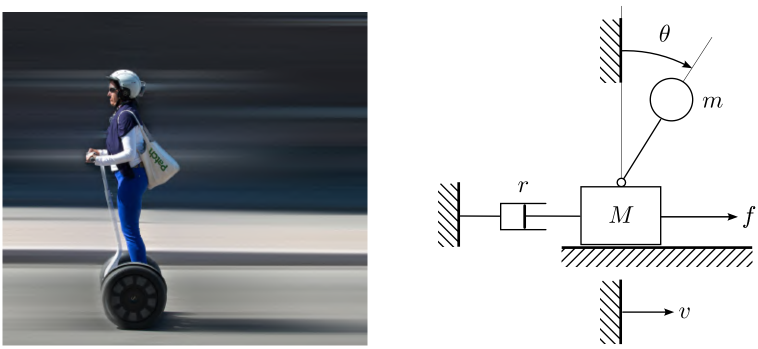
Notice that, since the matrix governing the evolution of the system obtain combining the Luenberger observer and the control action $u(t) = \boldsymbol{K}\hat{\boldsymbol{x}}(t)$ fed with the estimate $\hat{\boldsymbol{x}}(t)$ of the state provided by the observer itself, i.e. the matrix

$$\begin{bmatrix} \boldsymbol{A}+\boldsymbol{BK} & -\boldsymbol{BK} \\ \boldsymbol{0} & \boldsymbol{A}+\boldsymbol{LC} \end{bmatrix}$$

is block upper-triangular, its eigenvalues are those of matrix $\boldsymbol{A}+\boldsymbol{BK}$ and those of $\boldsymbol{A}+\boldsymbol{LC}$. We can then introduce the following result.

Property 21 (Separation principle)**.** *Provided that the strictly-proper LTI SISO system in* (11.3) *is both reachable and observable, it is possible to design an output feedback controller based on the combination of a Luenberger observer in* (11.5) *and the control law in* (11.4)*, for which the two gains* $\boldsymbol{K}$ *and* $\boldsymbol{L}$ *can be designed independently to arbitrarily assign all the eigenvalues of the closed-loop system.*

Example 68. *[Cart-pendulum] Consider the cart-pendulum system in Fig. 11.1(b). For*



(a) Segway (Creative Commons) (b) Model of the cart-pendulum system

Figure 11.1: A Segway vehicle and a cart-pendulum system.

small variations of angle θ around the upward equilibrium, i.e. for $\theta(t) \approx 0$, the equations governing the behaviour of the system are given in the following:

$$\begin{aligned} \frac{d}{dt}v(t) &= -\frac{r}{M+m}v(t) - \frac{mL}{M+m}\frac{d}{dt}\omega(t) + \frac{1}{M+m}f(t) \\ \frac{d}{dt}\theta(t) &= \omega(t) \\ \frac{d}{dt}\omega(t) &= \frac{1}{L}\frac{d}{dt}v(t) + \frac{g}{L}\theta(t) \end{aligned} \tag{11.10}$$

With the following set of parameters $M = 0.5$ kg, $m = 0.2$ kg, $r = 0.1$ Ns/m, $L = 0.3$ m, the dynamics governing the behaviour of the system can be characterised by the following matrices

$$\boldsymbol{A} = \begin{bmatrix} -0.1111 & -2.18 & 0 \\ 0 & 0 & 1 \\ -0.3704 & 25.4333 & 0 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 1.1111 \\ 0 \\ 3.7037 \end{bmatrix}$$

which is a reachable pair. Considering the angular displacement of the pendulum $\theta(t)$ as the measured output, the output equation is characterised by matrix

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

which, together with matrix $\boldsymbol{A}$, forms an observable pair.

A Luenberger observer designed so that the eigenvalues of $\boldsymbol{A}+\boldsymbol{L}\boldsymbol{C}$ are $s=-20$, $s=-30$, and $s=-40$ is characterised by parameter

$$\boldsymbol{L} = \begin{bmatrix} 64025.176 & -89.9 & -2615.445 \end{bmatrix}^T$$

In turn, the controller parameter $\boldsymbol{K}$ which place the eigenvalues of $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ in $s=-1$, $s=-2$, and $s=-3$ is the following one

$$\boldsymbol{K} = \begin{bmatrix} 0.2651 & -9.837 & -1.67 \end{bmatrix}$$

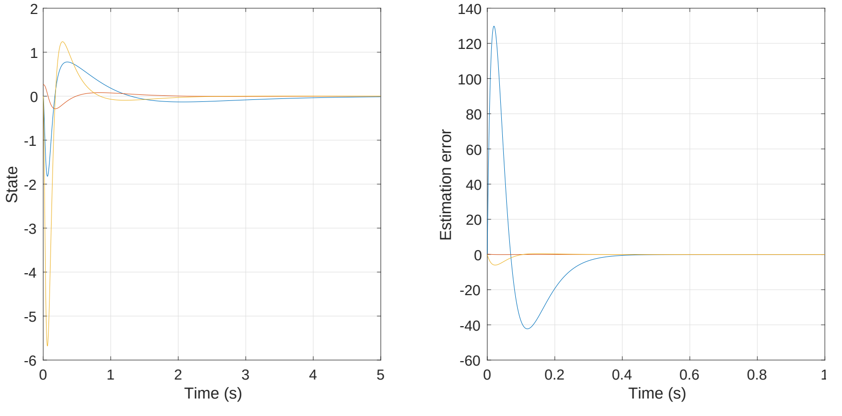


Figure 11.2: Evolution of the state vector $\boldsymbol{x}(t) = \begin{bmatrix} v(t) & \theta(t) & \omega(t) \end{bmatrix}^T$ (left, blue, red, and yellow, respectively) and of the estimation error $\boldsymbol{\varepsilon}(t)$.

Figure 11.2 reports the evolution of the state $\boldsymbol{x}(t)$ of the cart-pendulum system from initial condition $\boldsymbol{x}(0) = \begin{bmatrix} 0 & 15\,\text{deg} & 0 \end{bmatrix}^T$ together with the one of the estimation error $\varepsilon(t)$ (assuming $\hat{\boldsymbol{x}}(0) = \boldsymbol{0}$). As one can notice, the controller is capable of stabilising the pendulum in the upright position, $\theta \approx 0$, as well as to provide a null steady-state estimation error of its state. According to the design specification of closed-loop eigenvalues, the dynamics of the estimation error is way faster than that of the pendulum.

The overall block diagram of a complete state-space controller with a Luenberger observer $\mathcal{L}$ to govern the strictly proper reachable and observable LTI SISO system $G(s)$ is reported in Fig. 11.3.

The controller, which has $y(t)$ as input and $u(t)$ as output, is clearly a LTI SISO system. Its equations, which are those of the observer in (11.5) combined with one for the control law in (11.4), are given in the following:

$$\begin{aligned} \frac{d}{dt}\hat{\boldsymbol{x}}(t) &= (\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}+\boldsymbol{L}\boldsymbol{C})\,\hat{\boldsymbol{x}}(t) - \boldsymbol{L}y(t) \\ u(t) &= \boldsymbol{K}\hat{\boldsymbol{x}}(t) \end{aligned} \tag{11.11}$$

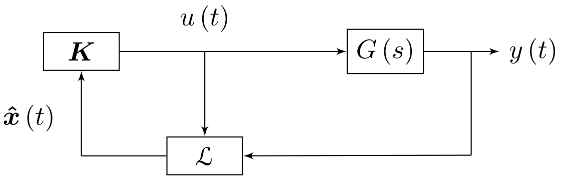


Figure 11.3: State-space controller with Luenberger observer $\mathcal{L}$.

which corresponds to the following transfer function

$$R(s) = -\boldsymbol{K}(s\boldsymbol{I} - \boldsymbol{A} - \boldsymbol{B}\boldsymbol{K} - \boldsymbol{L}\boldsymbol{C})^{-1}\boldsymbol{L}$$

Example 69. *[Cart-pendulum] Back to the previous example of the cart-pendulum system, the transfer function of the controller is the following one*

$$R(s) = -22226\frac{(s+4.912)(s-1.773)}{(s-9.509)(s^2+105.4s+4168)}$$

11.4 Zero-error regulation

Considering the output feedback controller developed in the previous Section, and looking at the corresponding closed-loop equations in (11.9), one can surely notice that the system has no inputs. Assuming asymptotic stability, the closed-loop system has the unique equilibrium $\bar{\boldsymbol{x}} = \bar{\boldsymbol{\varepsilon}} = \boldsymbol{0}$ and any trajectory of such a system will tend, sooner or later, to such an equilibrium point. As this property was somehow the objective of the cart-pendulum system, there are situations in which one would like to steer the state of the closed-loop system, or its output, to some other point rather than the origin. For example, in a possible application, one would like to track a certain reference $y^0(t)$ and provide a zero steady-state error whenever the reference is constant, i.e. $y^0(t) = \bar{y}, \forall t \geq 0$.

As we have seen in Chapters 7 and 8, in order to guarantee a zero steady-state error for a constant reference $y^0(t) = \bar{y}, \forall t \geq 0$, an integral action is needed within the controller. To handle this requirement, we **extend the state of the system** to be controlled with one additional equation, thus obtaining

$$\begin{aligned} \frac{d}{dt}\boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}u(t) \\ \frac{d}{dt}\xi(t) &= e(t) = y^0(t) - y(t) = y^0(t) - \boldsymbol{C}\boldsymbol{x}(t) \\ y(t) &= \boldsymbol{C}\boldsymbol{x}(t) \end{aligned} \tag{11.12}$$

Notice that, if the system is asymptotically stable, any equilibrium point will satisfy $e(t) = 0$. Before proceeding, a more compact form of the system at hands can be introduced

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} \boldsymbol{x}(t) \\ \xi(t) \end{bmatrix} &= \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ -\boldsymbol{C} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(t) \\ \xi(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{B} \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} \boldsymbol{0} \\ y^0(t) \end{bmatrix} \\ y(t) &= \begin{bmatrix} \boldsymbol{C} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(t) \\ \xi(t) \end{bmatrix} \end{aligned} \tag{11.13}$$

Then, provided that system (11.13) is both reachable and observable, we can proceed designing either a state feedback controller, or an output feedback controller combined with a Luenberger observer, depending on the accessibility of the state vector $\boldsymbol{x}(t)$. Notice that the state $\boldsymbol{x}_e(t)$ of the extended system (11.13) now comprises the state of the system $\boldsymbol{x}(t)$ and the state of the integrator $\xi(t)$ which is actually part of the controller, i.e.

$$\boldsymbol{x}_e(t) = \begin{bmatrix} \boldsymbol{x}(t)^T & \xi(t) \end{bmatrix}^T$$

The reachability of system (11.13) depends on the pair $(\boldsymbol{A}_e, \boldsymbol{B}_e)$ where

$$\boldsymbol{A}_e = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ -\boldsymbol{C} & 0 \end{bmatrix}, \boldsymbol{B}_e = \begin{bmatrix} \boldsymbol{B} \\ 0 \end{bmatrix}$$

Applying the Kalman's reachability test, after some trivial matrix algebra, we obtain the following reachability matrix for the extended system

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ -\boldsymbol{C} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{B} & \boldsymbol{AB} & \dots & \boldsymbol{A}^{n-1}\boldsymbol{B} \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ -\boldsymbol{C} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{K}_R \\ 1 & \boldsymbol{0}^T \end{bmatrix}$$

which is full-rank if and only if the pair $(\boldsymbol{A}, \boldsymbol{B})$ is reachable, i.e. if $\boldsymbol{K}_R$ is full-rank, and matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ -\boldsymbol{C} & 0 \end{bmatrix}$$

is non-singular.

By combining the equations of the integrator, i.e.

$$\frac{d}{dt}\xi(t) = y^0(t) - y(t)$$

those of the Luenberger observer (11.5) for the system (11.3), and the control action

$$u(t) = \boldsymbol{K}_e \begin{bmatrix} \hat{\boldsymbol{x}}(t) \\ \xi(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_x & K_\xi \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}}(t) \\ \xi(t) \end{bmatrix}$$

one obtains the equations describing the dynamics of the controller

$$\begin{aligned}\frac{d}{dt}\hat{\boldsymbol{x}}(t) &= \boldsymbol{A}\hat{\boldsymbol{x}}(t) + \boldsymbol{B}u(t) + \boldsymbol{L}\left(\boldsymbol{C}\hat{\boldsymbol{x}}(t) - y(t)\right)\\ \frac{d}{dt}\xi(t) &= y^0(t) - y(t)\\ u(t) &= \boldsymbol{K}_x\hat{\boldsymbol{x}}(t) + K_\xi\xi(t)\end{aligned} \tag{11.14}$$

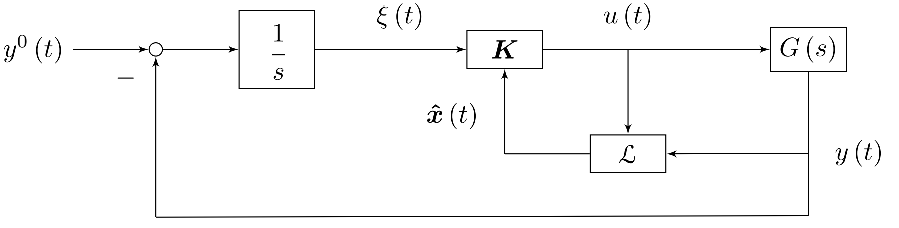


Figure 11.4: Zero-error regulation with Luenberger observer $\mathcal{L}$.

By rearranging the equations and substituting $y^0(t) - y(t)$ with $e(t)$ and similarly $y(t)$ with $y^0(t) - e(t)$, one obtains

$$\begin{aligned}\frac{d}{dt}\hat{\boldsymbol{x}}(t) &= (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{K}_x + \boldsymbol{L}\boldsymbol{C})\,\hat{\boldsymbol{x}}(t) + \boldsymbol{B}K_\xi\xi(t) - \boldsymbol{L}y^0(t) + \boldsymbol{L}e(t)\\ \frac{d}{dt}\xi(t) &= e(t)\\ u(t) &= \boldsymbol{K}_x\hat{\boldsymbol{x}}(t) + K_\xi\xi(t)\end{aligned} \tag{11.15}$$

which is a LTI system with one output $u(t)$ and two inputs, $y^0(t)$ and $e(t)$, respectively. The transfer function from $e(t)$ to $u(t)$, say $R(s)$, can be evaluated as

$$R(s) = \boldsymbol{K}_e\left(s\boldsymbol{I} - \boldsymbol{A}_R\right)^{-1}\boldsymbol{B}_R$$

while the transfer function from input $y^0(t)$ and output $u(t)$ is

$$C(s) = -\boldsymbol{K}_e\left(s\boldsymbol{I} - \boldsymbol{A}_R\right)^{-1}\boldsymbol{B}_C$$

where

$$\boldsymbol{A}_R = \begin{bmatrix}\boldsymbol{A} + \boldsymbol{B}\boldsymbol{K}_x + \boldsymbol{L}\boldsymbol{C} & \boldsymbol{B}K_\xi\\ \boldsymbol{0}^T & 0\end{bmatrix}, \boldsymbol{B}_R = \begin{bmatrix}\boldsymbol{L}\\ 1\end{bmatrix}, \boldsymbol{B}_C = \begin{bmatrix}-\boldsymbol{L}\\ 0\end{bmatrix}$$

Example 70. *[Mass-spring-damper] Consider once again the mass-spring-damper system of Fig. 2.2 with $m = 1$, $k = 10$, and $r = 0.1$ and assume again the output being the position of the mass, i.e. $y(t) = x_1(t)$. Matrix $\boldsymbol{\Sigma}$ results*

$$\boldsymbol{\Sigma} = \begin{bmatrix}0 & 1 & 0\\ -10 & -0.1 & 1\\ -1 & 0 & 0\end{bmatrix}$$

which is non-singular. We already verified the observability of the pair $(\boldsymbol{A}, \boldsymbol{C})$ *and designed the corresponding Luenberger observer having*

$$\boldsymbol{L} = \begin{bmatrix} -63.15 & -983.69 \end{bmatrix}^T$$

The reachability test of pair $(\boldsymbol{A}, \boldsymbol{B})$ *is left to the reader. Assuming the gain* $\boldsymbol{K}_e$ *has to be tuned to place the closed-loop eigenvalues in* $s = -3$, $s = -4$ *and* $s = -5$, *we obtain*

$$\boldsymbol{K}_e = \begin{bmatrix} -37 & -11.9 & 60 \end{bmatrix}$$

We finally obtain the transfer function of the controller:

$$R(s) = 14102 \frac{s^2 + 2.334s + 4.255}{s(s^2 + 75.15s + 1788)}$$

and the one of the feed-forward compensator

$$C(s) = -14042 \frac{s + 2.073}{(s^2 + 75.15s + 1788)}$$

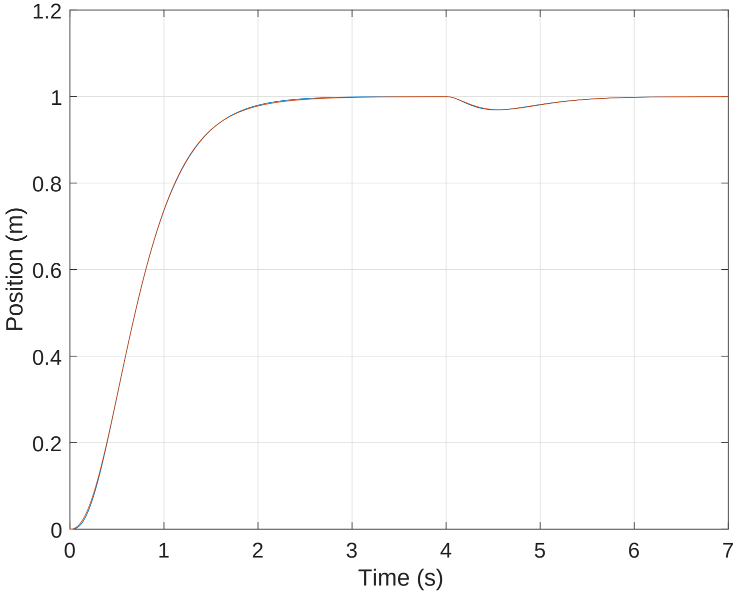


Figure 11.5: Response of the position of the mass-spring-damper system for $y^0(t) = 1, t \geq 0$ and a load disturbance $w(t) = -1, t \geq 5$ for controller $R(s)$ and feed-forward compensator $C(s)$ (blue), and for the approximating ISA PID (red).

Notice that the transfer function of the controller $R(s)$ *has a pole in the origin, two zeros, and two complex and conjugate poles at frequency* $\omega_n \approx 42.3\ rad/s$. *By substituting these two poles with a real one at the same frequency, the transfer function* $R(s)$ *can be approximated as*

$$R(s) \approx 7.8 \frac{s^2 + 2.334s + 4.255}{s(1 + s/42.3)}$$

which constitutes a PID controller with $K_p = 33.5461$, $T_i = 0.5483$, $T_d = 0.4285$, *and* $N = 18.1247$. *If the same approximation is performed on transfer function* $C(s)$, *we obtain*

$$C(s) \approx -16.274 \frac{1 + 0.4821s}{1 + 0.02364s}$$

which is the corresponding compensator for $b = 0.51487$ *and* $c = 0.48089$.

Figure 11.5 reports the time history of the position of the mass-spring-damper system in response to a step reference $y^0(t)$ *and a step load disturbance* $w(t)$ *applied after 4 s. As one can see, for both the two exogenous inputs, the system reaches a null steady-state error, as expected.*

In this Section, we have introduced a methodology to simultaneously design the transfer function of the feedback controller $R(s)$ and the one of the feed-forward compensator $C(s)$ based on the specification of the position of closed-loop eigenvalues. This can be seen as an alternative procedure to those based on the frequency response method in Chapter 8 and the one based on the root locus method of Chapter 10. While the root locus method seems not to have particular limitations on its applicability, the one based on the frequency response was only applicable to systems without poles with strictly positive real parts. In turn, guidelines for the design of the controller were easy to be derived for the latter case, while the root locus required some implicit reasoning on how the introduction of poles and/or zeros in $R(s)$ were capable of modifying the locus.

The method introduced in this Section somehow suffers from the same problem: it is somewhat difficult to predict how the positioning of eigenvalues for the closed-loop system are affecting, e.g., the attenuation of the effects of disturbances to variables of interest. Additionally, the requirement on the non-singularity of matrix $\boldsymbol{\Sigma}$ seems quite novel for the zero-error regulation problem and deserves further investigations.

Assume $\boldsymbol{A}$ to be an invertible matrix, i.e. $\det(\boldsymbol{A}) \neq 0$, then from the Schur's formula, the determinant of matrix $\boldsymbol{\Sigma}$ can be computed as

$$\det(\boldsymbol{\Sigma}) = \det(\boldsymbol{A}) \det\left(\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)$$

Notice that for SISO systems the quantity $\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}$ is actually a scalar quantity. Hence, provided that the system to be controlled does not have already a pole in the origin, i.e. $\det(\boldsymbol{A}) \neq 0$, the zero-error regulation is well-posed if and only if

$$\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \neq 0$$

This quantity is nothing else than the transfer function $G(s)$ evaluated for $s = 0$, in fact

$$G(s) = \boldsymbol{C}(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B}$$

Whenever $\det(\boldsymbol{A}) \neq 0$, assuming $\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \neq 0$ is equivalent to assume that the system to be controlled has **no zeros in the origin**. As we know, in fact, from the Kalman's decomposition, a cancellation between the pole in the origin to be added within the controller and one of the zeros in the origin of the system will make the corresponding dynamics of the system non reachable and/or non observable. If, from the one hand, this is preventing one to apply the design method introduced in this Chapter, from the other hand the stability of the part of the system which is not included in the feedback path will remain unaltered, hence surely not asymptotically stable.

Example 71. *[Cart pendulum] Consider again the cart-pendulum in Fig. 11.1(b), the corresponding matrix* $\mathbf{\Sigma}$ *is the following one*

$$\mathbf{\Sigma} = \begin{bmatrix} -0.1111 & -2.18 & 0 & 1.111 \\ 0 & 0 & 1 & 0 \\ -0.3704 & 25.4333 & 0 & 3.7037 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

is clearly singular, as one can notice looking at the last two columns or the last two rows. As a consequence, the reachability of the extended pair $(\boldsymbol{A}_e, \boldsymbol{B}_e)$ *cannot be granted, as discussed previously.*
Consistently, one would expect the transfer function $G(s)$ *of the system to have at least one zero in the origin, in fact*

$$G(s) = \frac{3.704s}{s^3 + 0.1111s^2 - 25.43s - 3.633}$$

Part III

Digital feedback control systems

Chapter 12

Dynamical systems in discrete-time

Consider a bus and the number of people presently on board as shown in Fig. 12.1. Clearly the number of people on the bus, say x, is a function of time, i.e. $x(t)$. On the other hand, its value changes only corresponding to bus stops and not continuously with time t. In order to better describe the dynamics of this

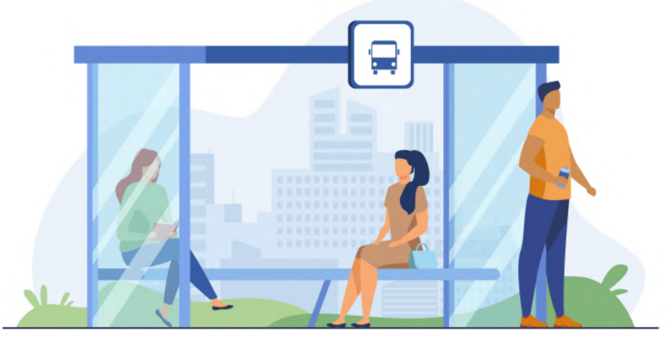


Figure 12.1: Bus station (designed by FreePik, www.freepik.com).

system one can consider the number of people on the bus prior to each station k, i.e. $x(k)$. Then, it is quite easy to model the dynamics of the system by observing that **after** station k the number of people on the bus, say $x(k+1)$ is equal to the number **before** station k, i.e. $x(k)$ plus the net number of people getting on board at station k (that is the difference between the ones getting on board and the ones getting of the bus), say $u(k)$, namely

$$x(k+1) = x(k) + u(k)$$

The model of the considered system has at least two fundamental differences with respect to the ones introduced in Chapter 2:

1. the state x evolves corresponding to discrete instants of time and not continuously;
2. the variable describing these time instants is no longer continuous and dense but discrete and countable.

12.1 Discrete-time systems

Let $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^n, \boldsymbol{u} \in \mathcal{U} \subseteq \mathbb{R}^{n_u}$ be the (column) vector representing the state and the (column) vector representing the $n_u \in \mathbb{Z}_{\geq 0}$ exogenous inputs, respectively, and $k \in \mathbb{N}$ be the variable representing the discrete time. Integer $n \in \mathbb{Z}_{>0}$ (i.e. the number of state variables) is called **order** of the system. Then the following system of equations represents the mathematical model of a (generic) discrete-time dynamical system:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} f_1(x_1(k), x_2(k), \ldots, x_n(k), u_1(k), u_2(k), u_{n_u}(k), k) \\ f_2(x_1(k), x_2(k), \ldots, x_n(k), u_1(k), u_2(k), u_{n_u}(k), k) \\ \vdots \\ f_n(x_1(k), x_2(k), \ldots, x_n(k), u_1(k), u_2(k), u_{n_u}(k), k) \end{bmatrix} \tag{12.1}$$

For the sake of compactness, we will make reference to the following (vectorial) form

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k), k) \tag{12.2}$$

with obvious meaning of symbols.

Definition 14 (Motion of the state). *Consider the generic discrete-time dynamical system in (12.2) and assume that state $\boldsymbol{x}(k)$ at discrete time $k = k_0$, i.e. $\boldsymbol{x}(k_0) \in \mathcal{X}$ is known. If the following problem*

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k), k) \\ \boldsymbol{x}(k_0) &\in \mathcal{X} \end{aligned} \tag{12.3}$$

*admits a unique solution for $k \geq k_0$, then this solution $\boldsymbol{x}(k)$ is called (state) **trajectory** associated to initial state $\boldsymbol{x}(k_0)$ and input $\boldsymbol{u}(k), k \geq k_0$ (if any).*

The concept of state trajectory is almost identical to the one introduced in Chapter 2. Similarly, one can introduce the property of a trajectory being stable as follows:

Definition 15 (Stability of a trajectory). *A trajectory $\boldsymbol{x}(k), k \geq k_0$ generated from initial state $\boldsymbol{x}(k_0)$ and input $\boldsymbol{u}(k), k \geq k_0$ is called **stable** if $\forall \varepsilon > 0$ that is "small enough", there exists $\delta > 0$ such that if $\|\boldsymbol{x}(k_0) - \tilde{\boldsymbol{x}}(0)\| < \delta$, then $\forall k \geq k_0 : \|\boldsymbol{x}(k) - \tilde{\boldsymbol{x}}(k)\| < \varepsilon$.*

The definition of stability is also identical to the one introduced in Chapter 2 for trajectories generated by continuous-time dynamical systems, provided that t is substituted with k. Unstable and asymptotically stable trajectories are defined in similar ways.

12.1.1 Equilibria

Similarly to the continuous-time case, equilibria are particular state trajectories characterised by a constant value of the state vector, i.e.

$$\boldsymbol{x}(k+1) = \boldsymbol{x}(k), \forall k$$

Formally

Definition 16 (Equilibrium (constant trajectory))**.** *State trajectory* $\boldsymbol{x}(k) = \bar{\boldsymbol{x}}, \forall k \geq k_0$ *is called* ***equilibrium*** *if* $\exists \bar{\boldsymbol{u}} : \boldsymbol{f}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}, k) = \bar{\boldsymbol{x}}, \forall k \geq k_0$.

Example 72. *Assume one wants to solve for $z \in \mathbb{R}$ the following non-linear equation*

$$z - f(z) = 0$$

A way to find the solution of such a non-linear equation is to set $x(0) = x_0$ for some value of x_0 and iterate according to the following law

$$x(k+1) = f(x(k))$$

which represents a first-order non-linear discrete-time dynamical system, with k the index identifying the iterations. Notice that the equilibria for this system are the roots of the equation, in fact

$$\bar{x} = f(\bar{x})$$

If, for example, to solve the non-linear equation characterised by $f(z) = -z^3$, i.e. $z+z^3 = 0$, starting from $x_0 = -0.75$, one obtains

$$x(1) = f(x(0)) = f(x_0) = -0.75^3 \approx -0.4219$$
$$x(2) = f(x(1)) = -(-0.4219)^3 \approx 0.0751$$
$$x(3) = f(x(2)) = -\left(0.0751^3\right) \approx -0.0004233$$

which clearly tends to $\bar{x} = 0$ for k large enough, which is one solution of $z + z^3 = 0$.

12.1.2 Systems with outputs

As in the case of continuous-time, discrete-time systems might be characterised by output variables $\boldsymbol{y} \in \mathbb{R}^{n_y}$ and how these variables related to the state vector, the input vector, and the discrete time as follows

$$\begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_{n_y}(k) \end{bmatrix} = \begin{bmatrix} h_1(x_1(k), x_2(k), \ldots, x_n(k), u_1(k), u_2(k), u_{n_u}(k), k) \\ h_2(x_1(k), x_2(k), \ldots, x_n(k), u_1(k), u_2(k), u_{n_u}(k), k) \\ \vdots \\ h_{n_y}(x_1(k), x_2(k), \ldots, x_n(k), u_1(k), u_2(k), u_{n_u}(k), k) \end{bmatrix} \tag{12.4}$$

or, in short:

$$\boldsymbol{y}(k) = \boldsymbol{h}(\boldsymbol{x}(k), \boldsymbol{u}(k), k) \tag{12.5}$$

Consider for example a computing system constituted of a CPU, a memory, a keyboard, and a monitor (see Fig. 12.2). At every clock tick $k \in \mathbb{N}$, the CPU

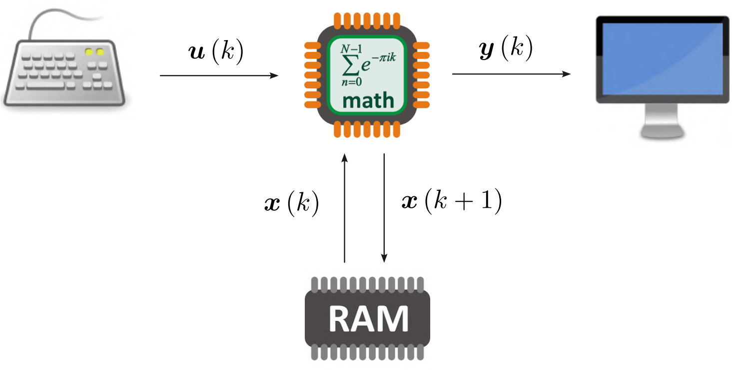


Figure 12.2: A computing system.

unit reads the state of the memory, say $\boldsymbol{x}(k)$, and, based on possible input keys $\boldsymbol{u}(k)$ updates the memory itself based on some algebraic operations, i.e.

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k))$$

Consistently, the CPU produces some output $\boldsymbol{y}(k)$ to be displayed on the monitor, i.e.

$$\boldsymbol{y}(k) = \boldsymbol{h}(\boldsymbol{x}(k), \boldsymbol{u}(k))$$

It is then easy to realise that the behaviour of a computing system evolves corresponding to the discrete clock ticks of the CPU and can be then represented as a discrete-time dynamical system.

12.2 Discrete-time LTI systems in state-space

Discrete-time LTI systems constitute a particular class of discrete-time dynamical systems characterised by linear functions $\boldsymbol{f}$ and $\boldsymbol{h}$ with no explicit dependency on time k.

Definition 17 (LTI system)**.** *A discrete-time LTI system is a particular discrete-time dynamical system which can be written in the following form*

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k) \\ \boldsymbol{y}(k) &= \boldsymbol{C}\boldsymbol{x}(k) + \boldsymbol{D}\boldsymbol{u}(k) \end{aligned} \tag{12.6}$$

12.2.1 State trajectories for discrete-time LTI systems and some insights on their stability

Similar to the case of continuous-time LTI systems, for discrete-time LTI systems in the form (12.6) there is a formula to explicitly compute its state trajectories.

Theorem 9. *The following IVP*

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k), \boldsymbol{x}(0) \in \mathcal{X}$$

is solved by the following trajectory

$$\boldsymbol{x}(k) = \boldsymbol{A}^k \boldsymbol{x}(0) + \sum_{i=0}^{k-1} \boldsymbol{A}^{k-i-1}\boldsymbol{B}\boldsymbol{u}(i), k \geq 0 \tag{12.7}$$

Proof. Left to the reader. □

Notice that, also in the case of discrete-time LTI systems, state trajectories are formed by a **free response** which depends on the initial state $\boldsymbol{x}(0)$ only, and by a **forced response**, which in turn depends on the input $\boldsymbol{u}(k)$ only. Moreover, the stability of a trajectory generated by a discrete-time LTI system can be determined by the its free response, only. Consider in fact a trajectory $\boldsymbol{x}(k), k \geq 0$ is originated by applying input $\boldsymbol{u}(k)$ from the initial state $\boldsymbol{x}(0)$. If the same input is applied from a perturbed initial condition $\tilde{\boldsymbol{x}}(0)$, the perturbed state trajectory can be now evaluated as:

$$\tilde{\boldsymbol{x}}(k) = \boldsymbol{A}^k \tilde{\boldsymbol{x}}(0) + \sum_{i=0}^{k-1} \boldsymbol{A}^{k-i-1}\boldsymbol{B}\boldsymbol{u}(i), k \geq 0$$

As stability of $\boldsymbol{x}(k)$ depends on the possibility to obtain a bounded perturbation on the state trajectory, i.e. $\|\boldsymbol{x}(k) - \tilde{\boldsymbol{x}}(k)\| < \varepsilon$ provided that the perturbation on the initial condition is bounded, i.e. $\|\boldsymbol{x}(0) - \tilde{\boldsymbol{x}}(0)\| < \delta$, one is actually requiring

$$\|\boldsymbol{x}(k) - \tilde{\boldsymbol{x}}(k)\| = \left\|\boldsymbol{A}^k (\boldsymbol{x}(0) - \tilde{\boldsymbol{x}}(0))\right\| < \varepsilon$$

12.2.2 Stability of discrete-time LTI systems

Similarly to the continuous-time case, stability of state trajectories of a discrete-time LTI, and stability of the system as a whole, can be addressed by analysing whether $\left\|\boldsymbol{A}^k (\boldsymbol{x}(0) - \tilde{\boldsymbol{x}}(0))\right\|$ can be made bounded by requiring a proper bound on the perturbation on the initial condition. It should be clear that stability should be strongly related with term $\boldsymbol{A}^k$ and therefore with matrix $\boldsymbol{A}$ which characterises the system.

Stability and eigenvalues

Similarly to the case of continuous-time LTI systems, stability can be addressed based on the analysis of the eigenvalues of matrix $\boldsymbol{A}$.

Theorem 10 (Stability analysis of discrete-time LTI systems)**.** *Depending on matrix $\boldsymbol{A}$, system* (12.6) *is*

- *asymptotically stable if and only if all the eigenvalues of $\boldsymbol{A}$ have magnitude strictly less than one;*
- *stable (but not asymptotically) if and only if all the eigenvalues of $\boldsymbol{A}$ have magnitude at most equal to one, and those unit magnitude are simple roots;*
- *unstable if and only if either at least one of the eigenvalues of $\boldsymbol{A}$ has magnitude higher than one, or all the eigenvalues of $\boldsymbol{A}$ have magnitude not higher than one, and those with unit magnitude are not simple roots.*

Stability and characteristic polynomial

Differently from the continuous-time case, asymptotic stability requires all the eigenvalues to have magnitude strictly lower than one. Though a method similar to the Routh-Hurwitz criterion exists, we are looking for a way to apply the Routh-Hurwitz criterion itself. The criterion cannot be clearly applied as is (as it will check whether the eigenvalues of matrix $\boldsymbol{A}$ are within the left half-plane, which is of no particular interest for discrete-time LTI systems). We then need to find a way to transform the characteristic polynomial of matrix $\boldsymbol{A}$ in such a way that every eigenvalue with magnitude strictly less than one is mapped into an eigenvalue with strictly negative real part. Such a transformation exists and is called **bilinear transform**.

Property 22 (Bilinear transform)**.** *Consider the characteristic equation*

$$\pi(\lambda) = 0$$

then its roots have magnitude strictly lower than one if and only if the roots of

$$\pi\left(\frac{1+\bar{\lambda}}{1-\bar{\lambda}}\right) = 0$$

are all with strictly negative real parts.

In short, the transform just introduced (which is called **bilinear transform**) is mapping the interior of the unit circle of the complex plane, i.e. $|\lambda| < 1$, into the open left half-space of the complex plane, i.e. $\mathbb{R}e\left[\bar{\lambda}\right] < 0$. In fact, let

$$\bar{\lambda} = \alpha + j\beta$$

then

$$|\lambda|^2 = \left|\frac{1+\bar{\lambda}}{1-\bar{\lambda}}\right|^2 = \frac{|1+\alpha+j\beta|^2}{|1-\alpha-j\beta|^2} = \frac{1+\alpha^2+2\alpha+\beta^2}{1+\alpha^2-2\alpha+\beta^2} < 1 \Leftrightarrow \alpha < 0$$

The asymptotic stability property of a discrete-time LTI system can be then verified still resorting to the Routh-Hurwitz criterion, provided that the bilinear transform is first applied to the characteristic polynomial of $\boldsymbol{A}$ which characterises the system.

Example 73. *[Third order system] A discrete-time LTI system is characterised by matrix*

$$\boldsymbol{A} = \begin{bmatrix} 7/30 & 0 & 2/15 \\ 4/15 & -1/10 & -8/15 \\ 1/15 & 0 & 1/6 \end{bmatrix}$$

which corresponds to the following characteristic polynomial

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^3 - 3/10\lambda^2 - 1/100\lambda + 3/1000$$

By numerically solving the characteristic equation one obtains the following eigenvalues

$$\lambda \in \{-1/10, 3/10, 1/10\}$$

clearly showing that the corresponding discrete-time LTI system is asymptotically stable. On the other hand, applying the bilinear transform one obtains

$$\left(\frac{1+\bar{\lambda}}{1-\bar{\lambda}}\right)^3 - 3/10\left(\frac{1+\bar{\lambda}}{1-\bar{\lambda}}\right)^2 - 1/100\left(\frac{1+\bar{\lambda}}{1-\bar{\lambda}}\right) + 3/1000$$

which has the same roots of the same multiplied by $\left(1-\bar{\lambda}\right)^3$, *i.e.*

$$\left(1+\bar{\lambda}\right)^3 - 3/10\left(1+\bar{\lambda}\right)^2\left(1-\bar{\lambda}\right) - 1/100\left(1+\bar{\lambda}\right)\left(1-\bar{\lambda}\right)^2 + 3/1000\left(1-\bar{\lambda}\right)^3$$

Finally, the Routh-Hurwitz criterion can be applied to the following characteristic equation

$$1287\bar{\lambda}^3 + 3319\bar{\lambda}^2 + 2701\bar{\lambda} + 693 = 0 \tag{12.8}$$

1287	2701	0
3319	693	0
2432.28	0	
693		

Since the first column has all entries with the same positive sign, the eigenvalues of (12.8) *have all strictly negative real parts which, in turn, implies that matrix* $\boldsymbol{A}$ *has all eigenvalues with magnitudes strictly lower than one.*

12.2.3 Asymptotically stable discrete-time LTI systems with inputs

As done for continuous-time LTI systems in Chapter 2, we can introduce few properties for asymptotically stable discrete-time LTI systems.

Property 23. *Consider the discrete-time LTI system in* (12.6) *and assume it is asymptotically stabile, then*

1. *for a constant input* $\boldsymbol{u}(k) = \bar{\boldsymbol{u}}, \forall k \geq 0$, *the equilibrium is unique and characterised by*
$$\bar{\boldsymbol{x}} = (\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{B}\bar{\boldsymbol{u}}$$

2. *after a sufficiently long time, the trajectory of the state depends only on the input, as the free response vanishes;*

3. *if the input* $\boldsymbol{u}(k)$ *tends to a constant value* $\bar{\boldsymbol{u}}$, *then the trajectory of the state tends to the corresponding equilibrium* $(\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{B}\bar{\boldsymbol{u}}$.

12.3 Transfer function of discrete-time LTI SISO systems

Similar to the continuous-time case, discrete-time LTI systems can be represented in the frequency domain. Consider the following discrete-time LTI SISO system

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}u(k) \\ y(k) &= \boldsymbol{C}\boldsymbol{x}(k) + \boldsymbol{D}u(k) \\ \boldsymbol{x}(0) &= \boldsymbol{x}_0 \end{aligned} \tag{12.9}$$

and apply the Z-transform to $u(k)$, $\boldsymbol{x}(k)$, and $y(k)$

$$\mathcal{Z}[u(k)] = U(z), \mathcal{Z}[\boldsymbol{x}(k)] = \boldsymbol{X}(z), \mathcal{Z}[y(k)] = Y(z)$$

and

$$\mathcal{Z}[\boldsymbol{x}(k+1)] = z(\boldsymbol{X}(z) - \boldsymbol{x}(0))$$

By substituting the latter in the equation of the system, one obtains

$$z(\boldsymbol{X}(z) - \boldsymbol{x}(0)) = \boldsymbol{A}\boldsymbol{X}(z) + \boldsymbol{B}U(z)$$

$$Y(z) = \boldsymbol{C}\boldsymbol{X}(z) + \boldsymbol{D}U(z)$$

Rearranging the two previous equations, one obtains

$$Y(z) = \left[\boldsymbol{C}(z\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{B} + \boldsymbol{D}\right] U(z) + \boldsymbol{C}(z\boldsymbol{I} - \boldsymbol{A})^{-1} z\boldsymbol{x}(0)$$

where the scalar quantity $G(z) = \boldsymbol{C}(z\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{B} + \boldsymbol{D}$ is called **transfer function** and represents the relationship between the Z-transform of the input $U(z)$ and the Z-transform of the forced response of the output $Y^f(z)$, so that

$$Y^f(z) = G(z)U(z)$$

Definition 18. *Given a discrete-time LTI SISO system and its transfer function $G(z)$:*

1. *a root of the numerator is called **zero**;*
2. *a root of the denominator is call **pole**;*
3. *the difference between the degree of the denominator and the degree of the numerator $\nu \geq 0$ is called **relative degree**;*
4. *if the relative degree is positive ($\nu > 0$), which happens whenever $D = 0$, the transfer function is called **strictly proper**, otherwise proper;*
5. *the number $g \in \mathbb{Z}$ of poles in $z = 1$ is called **type**;*
6. *quantity $\mu = \lim_{z \to 1} (z-1)^g G(z)$ is called **generalised gain**.*

Chapter 13

Digital control systems

The objective of this Chapter is twofold: identify how a discrete-time transfer function $R_d(z)$ can be derived to mimic the behaviour of $R(s)$, and understand how $R_d(z)$ can be translated into a set of instructions to be executed by a CPU. Transfer function $R(s)$ is assumed to be available as an output of one of the design strategies discussed in Chapters 8-11.

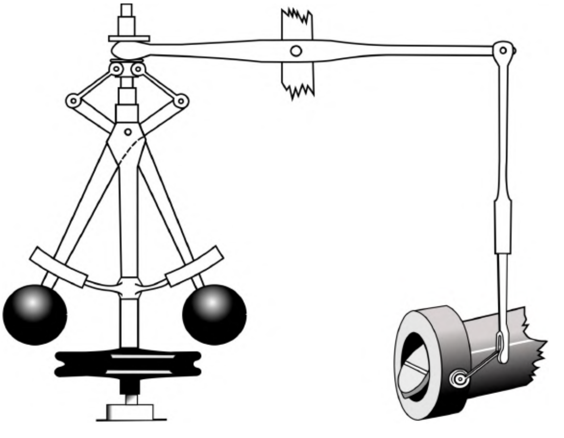


Figure 13.1: Watt's (or centrifugal) governor (Creative Commons).

13.1 Motivations

In the past, feedback control systems were characterised by analogue devices playing the role of the controller. In other words, given a system to be controlled, the transfer function of the controller $R(s)$ were derived based on one of the methods introduced earlier in this book and then a suitable electrical circuit or a mechanical device was designed to have an input-output relationship similar to the one in $R(s)$.

For example, Fig. 13.1 reports a centrifugal governor adopted by James Watt in the late 1700's to control the speed of a steam engine. The governor is connected to a throttle valve that regulates the flow of the steam into the engine. As the speed increases, the speed of the main axis of the governor also increases, and the two masses move outwards and upwards, reducing the aperture of a throttle valve, and eventually the speed of the engine.

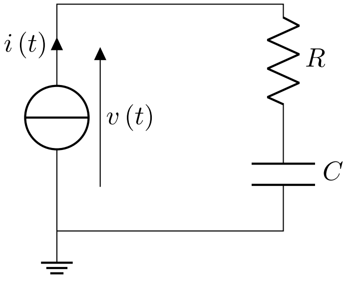


Figure 13.2: Analogue implementation of a proportional-integral controller.

As another example, consider the electrical circuit in Fig. 13.2. Letting $i(t)$, the current flowing in the circuit, the voltage drop on the capacitor is related to such a current as

$$C\frac{d}{dt}v_C(t) = i(t)$$

Moreover, the voltage drop on the resistance is

$$v_R(t) = Ri(t)$$

Therefore, if the current $i(t)$ represents the error $e(t)$ between a reference and the corresponding measured variable, the voltage $v(t)$ provided by the current supplier in Fig. 13.2 can be computed as

$$v(t) = v_R(t) + v_C(t) = Ri(t) + \frac{1}{C}\int_0^t i(\tau)\,d\tau$$

which clearly represents the control law of a proportional-integral controller having R and $1/C$ as proportional and integral gains, respectively.

Modern feedback control systems no longer adopt electrical circuits or mechanisms to physically implement the controller. They rather rely on computing systems whose behaviour is **designed to mimic** the one of transfer function $R(s)$. The word *mimic* has been introduced because, as seen in Chapter 12 and in particular in Fig. 12.2, a computing system has to be represented in discrete-time system and, if the algorithm is linear, in terms of a discrete-time transfer function $R_d(z)$.

13.2 Computer controlled systems

The complete block diagram of a computer-controlled system is reported in Fig. 13.3. Differently from the one in Fig. 7.1, the controller is now represented by the discrete-time transfer function $R_d(z)$ and emits the discrete-time signal $u_d(k)$. Two blocks have been added: the **digital to analogue** converter (D/A) which transforms the discrete-time signal $u_d(k)$ into the continuous-time one $u(t)$ and the **analogue to digital** converter (A/D) which plays the opposite role and converts a continuous-time signal into a discrete-time one.

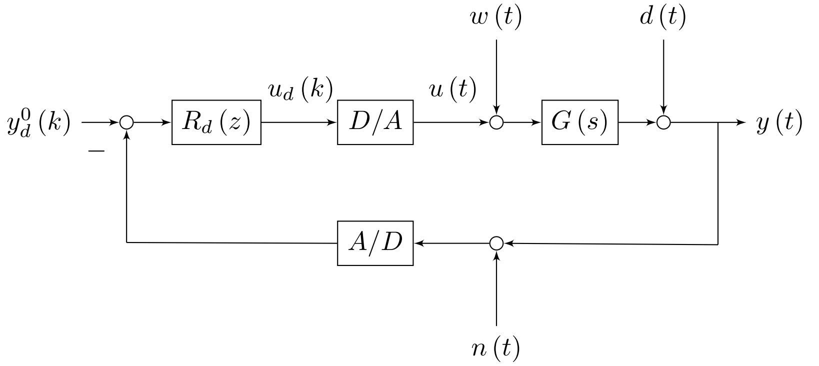


Figure 13.3: A negative feedback computer controlled system.

In the remainder of this Chapter, we will model the behaviour of such subsystems for their possible impact in the behaviour of the overall control system. Technological insights on how these functionalities will be implemented in hardware will be given in Chapter 14.

13.2.1 From continuous-time to discrete-time

The role of the A/D converter is primarily the conversion from a continuous-time signal, say $v(t)$, to a discrete-time one, $v_d(k)$. This functionality is simply realised by **sampling** the signal at a constant rate, i.e. the **sampling frequency** $f_s = 1/T_s$, where T_s is called **sampling time**. The relationship between $v(t)$ and $v_d(k)$ is therefore given in the following

$$v_d(k) = v(kT_s), k \in \mathbb{N}$$

An example of the sampling procedure is reported in Fig. 13.4. Notice from Fig. 13.4 that the samples $v_d(k)$ are well-representative of the original continuous-time signal $v(t)$ and a simple linear interpolation of the samples will allow one to reasonably reconstruct $v(t)$.

When $\sin(2\pi t)$ is added to $v(t)$, the situation results dramatically different, as shown in Fig. 13.5. From the one hand, the simple linear interpolation of

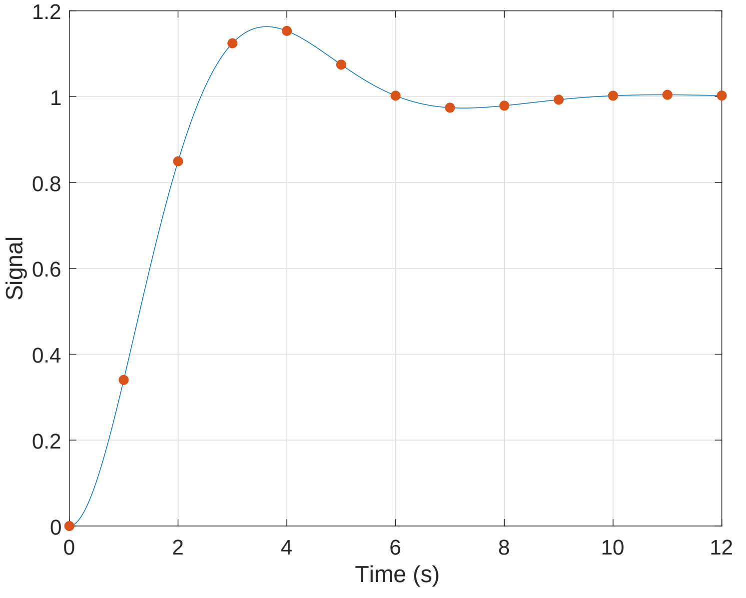


Figure 13.4: Continuous-time signal $v(t)$ (blue) and corresponding samples $v_d(k)$ (red dots) for $T_s = 1$ seconds.

samples is not capable of accurately reconstruct the original signal. On the other hand, since the samples of Fig. 13.4 and the ones in Fig. 13.5 are exactly the same, it is actually impossible to distinguish which of the two signals the samples have been extracted from.
As a conclusion, given a set of samples $v_d(k)$, there are infinitely many signals $v(t)$ such that $v_d(k) = v(kT_s)$.

Theorem 11 (Nyquist–Shannon sampling theorem)**.** *If the continuous-time signal to be sampled $v(t)$ has a limited bandwidth, say $\omega^{\max}$, and the sampling time T_s is such that $\omega^{\max} < \dfrac{\pi}{T_s}$ (π/T_s being the so-called **Nyquist frequency**, ω_N), then the samples are "fully informative" of the given signal.*

For the example before, the frequency components of the continuous-time signal $v(t)$ are represented in Fig. 13.6. As one can see, the spectral components of the continuous-time signal are negligible after $\omega^{\max} \approx 0.5\ rad/s$ which is surely smaller than the Nyquist frequency $\omega_N = \pi/T_s = \pi \approx 3.14\ rad/s$. In turn, when $\sin(2\pi t)$ has been superimposed to $v(t)$, $\omega^{\max}$ would result equal to $2\pi \approx 6.28\ rad/s$, hence violating the Nyquist-Shannon theorem.

The Nyquist-Shannon theorem is relevant in many applications of engineering. For example, when an audio track has to be sampled and digitalised, one has to typically consider audible frequencies, ranging from 20 Hz to 20 kHz. It

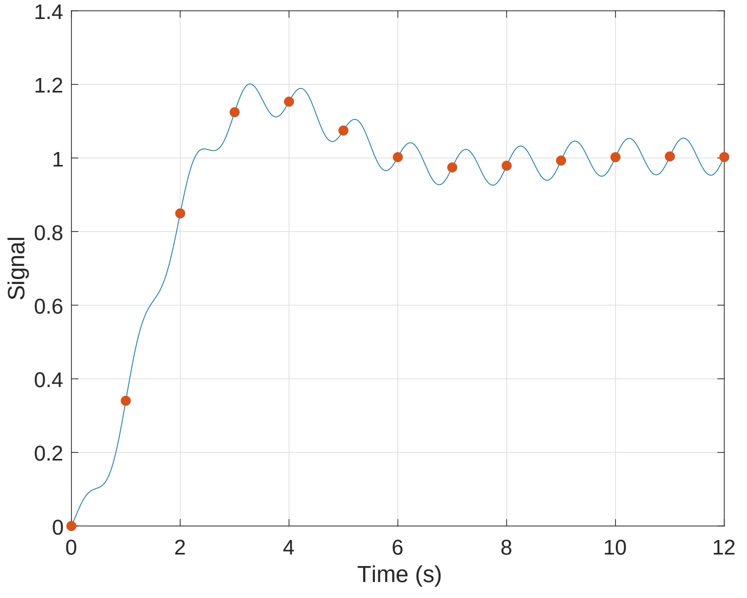


Figure 13.5: Continuous-time signal $v\left(t\right)+\sin\left(2\pi t\right)$ (blue) and corresponding samples $v_d\left(k\right)$ (red dots) for $T_s = 1$ seconds.

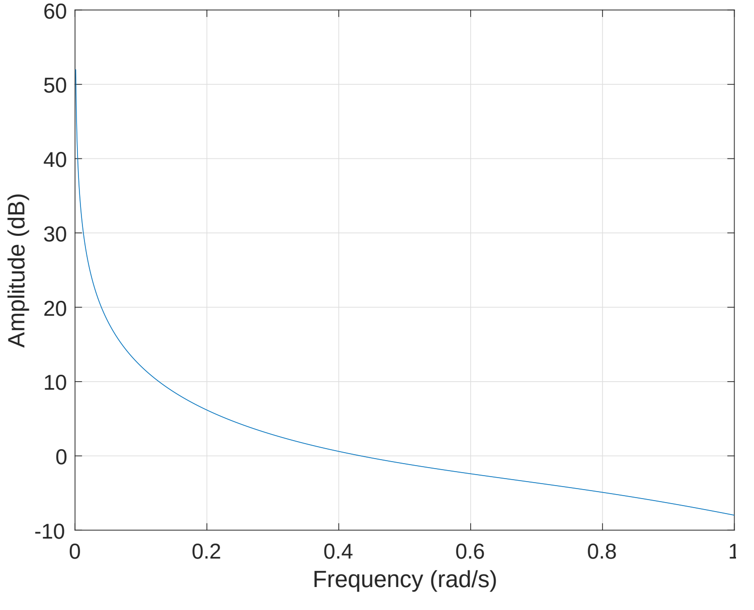


Figure 13.6: Spectral components of the signal $v\left(t\right)$ of Fig. 13.4.

should not be surprising that digital music in MP3 format (formally, MPEG-1 Audio Layer-3) adopts a sampling rate of $f_s = 48\ kHz$ which is slightly higher than the double of the maximum audible frequency.
As per digital control systems, the Nyquist-Shannon theorem can be clearly applied to select a proper sampling time T_s, once a value for $\omega^{\max}$ can be guaranteed. With reference to the block diagram of Fig. 13.3, the signal to be sampled is constituted by measured output $y(t)$ corrupted with the measurement noise $n(t)$. As per the reasoning reported in Chapter 8 it is reasonable to assume the bandwidth of signal $y(t)$ to be limited to up to a value that can be approximated with the **crossover frequency** ω_c. The measurement noise $n(t)$, in turn, deserves an additional analysis. Since an upper bound of the bandwidth of the noise characterising the measured variable will require to have very-high sampling frequencies, a low-pass filter is typically adopted on the measured value $y(t) + n(t)$. The cut-off frequency ω_f of such a low-pass filter which has the following transfer function

$$H(s) = \frac{1}{1 + s/\omega_f}$$

has to be selected sufficiently higher than the crossover frequency ω_c, but lower than π/T_s, i.e.

$$\omega_c \ll \omega_f \ll \omega_N = \frac{\pi}{T_s}$$

This way, while the components of $y(t)$ are preserved, those of the measurement noise $n(t)$ after ω_f are reduced before the measured variable can be sampled. Notice that, since $\omega_c \ll \omega_f$, the filter does not significantly modify the designed crossover frequency, which remains approximately equal to ω_c. On the other hand, the following **reduction on the phase margin**

$$\Delta\psi_m = -\text{atan}\left(\frac{\omega_c}{\omega_f}\right)$$

is be expected.

In extreme summary, the Nyquist frequency ω_N has to be selected sufficiently higher than the crossover frequency ω_c. Consequently the sampling time T_s has to satisfy the following inequality

$$T_s \ll \frac{\pi}{\omega_c}$$

Whilst the Nyquist-Shannon theorem puts no lower bounds on the sampling time T_s, it is important to realise that the smaller T_s, the more expensive and performative the hardware implementing the controller.

13.2.2 From discrete-time to continuous-time

With reference to Fig. 13.3, the role of the D/A converter is to transform the discrete-time signal $u_d(k)$ into a continuous-time one, $u(t)$. In the previous Section, we have seen that a continuous-time signal can be obtained from samples using some interpolating function. Let us consider the simple case of piece-wise constant function in which the value of $u_d(k)$ is maintained constant until the next discrete-time instant $k+1$, i.e.

$$u(t) = u_d(k), \forall t \in [kT_s, (k+1)T_s)$$

Figure 13.7 reports an example of the application of this D/A conversion.

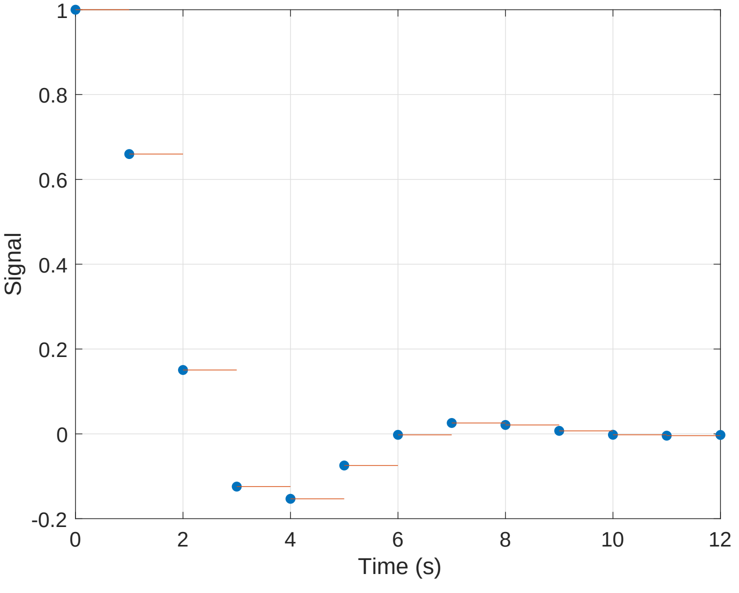


Figure 13.7: Discrete-time signal (blue dots) and its piece-wise constant continuous-time approximation (red).

The conversion method just introduced is called **zero order holder** (ZOH), as it maintains constant along the sampling interval the value of the 0-th order derivative (i.e. the value itself). In contrast, higher-older holders will maintain constant the corresponding derivative of the reconstructed signal. For example, first-order holder will be characterised by a piece-wise linear signal $u(t)$.

13.2.3 Effects of the A/D and D/A conversions on the control system

The two elements introduced in the previous Sections, i.e. the A/D and the D/A, are part of the computer controlled control loop of Fig. 13.3. The aim of

the next discussion is to provide the reader some insight regarding the consequences on the control system of the presence of these two blocks, in contrast with the ideal control system of Fig. 7.1 based on which the controller has been designed. For example, we have already seen that the sampler is typically implemented in conjunction with a low-pass filter and that this filter can somehow decrease the phase margin.
With reference again to Fig. 13.3, it seems interesting to address how the continuous-time signal feeding the digital controller is processed and possibly modified before feeding the system $G(s)$. Apart from the manipulations performed by the digital controller $R_d(z)$, the signal $y(t)$ has to traverse first the A/D (sampler) and then the D/A converter (ZOH).

$$v(t) \longrightarrow \boxed{\text{sampler}} \longrightarrow \boxed{ZOH} \longrightarrow \tilde{v}(t)$$

Figure 13.8: Series of a sampler (A/D) and a zero order holder (D/A).

Figure 13.9 reports the three signals $v(t)$, $v_d(k)$ (the output of the sampler), and $\tilde{v}(t)$ of Fig. 13.8 for a sampling time of $T_s = 1\ s$. As we have already discussed, the sampling time T_s satisfies the Nyquist-Shannon theorem and can be then considered adequate to sample signal $v(t)$. On the other hand, still referring to Fig. 13.8, the series composed by the sampler and the ZOH seem to introduce a certain delay. Based on some advanced arguments, that we skip for brevity, such a delay can be estimated as one half of the sampling time T_s.

Proposition 2. *The series of a sampler with sampling time T_s and a zero order holder is equivalent to a delay of $T_s/2$. Therefore, the relationship between the Laplace transforms of $v(t)$ and $\tilde{v}(t)$ can be regarded as*

$$\tilde{V}(s) = e^{-s\frac{T_s}{2}} V(s)$$

In summary, given a feedback control system with a crossover frequency ω_c, the sampling time T_s has to be selected so that

$$T_s \ll \frac{\pi}{\omega_c}$$

In addition, the reduction on the phase margin due to both the low-pass filter positioned before the sampler and the equivalent delay of $T_s/2$ introduced by the combination of the sampler and the ZOH has to be accounted. In particular, the presence of the sampler and of the ZOH will cause a **reduction of the phase margin** of

$$\Delta\psi_m = \omega_c \frac{T_s}{2} \ll \frac{\pi}{2}$$

If, for example, the sampling time T_s is selected as

$$T_s = \frac{\pi}{\alpha\omega_c}, \alpha \gg 1$$

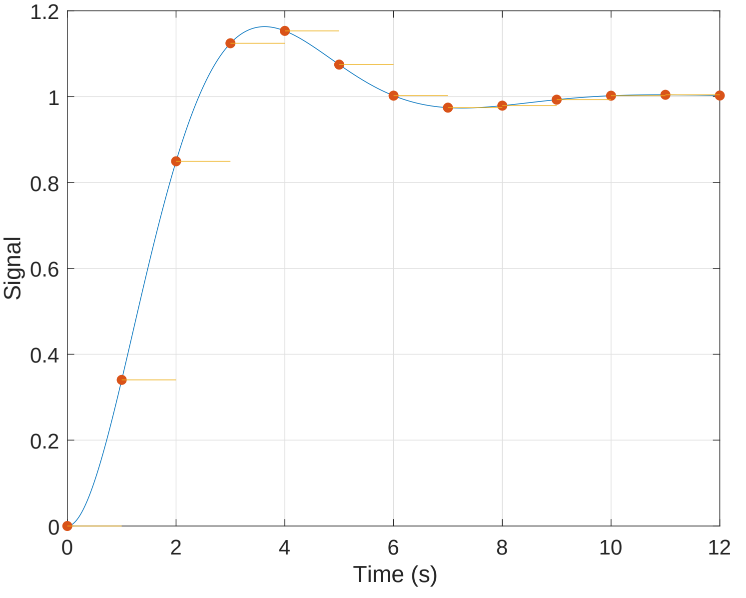


Figure 13.9: A continuous-time signal $v(t)$ (blue), the corresponding samples (red dots) for $T_s = 1$, and the signal $\tilde{v}(t)$ (yellow) reconstructed through a ZOH.

the corresponding reduction in the phase margin will result

$$\Delta\psi_m = \omega_c \frac{\pi}{2\alpha\omega_c} = \frac{\pi}{2\alpha} \ll \frac{\pi}{2}$$

13.3 Digital control law

Once the sampling T_s time has been carefully selected to satisfy the Nyquist-Shannon theorem as well as to limit the decrease in the phase margin ψ_m, a way to derive the transfer function of the discrete-time controller $R_d(z)$ out from $R(s)$ is needed.
Observing that z^{-1} of the Z-transform corresponds to a delay of T_s seconds in the Laplace domain, one can simply state that $z^{-1} = e^{-sT_s}$, from which

$$s = \frac{1}{T_s} \log_e(z)$$

and therefore

$$R_d(z) = R\left(\frac{1}{T_s} \log_e(z)\right)$$

which is however not a discrete-time transfer function as it is not represented by the ratio of two polynomials in variable z.

This issue can be solved by evaluating a rational approximation of the $\log_e$ function, i.e.

$$\log_e(z) \approx 2\frac{z-1}{z+1}$$

Therefore, the discrete-time transfer function of the controller $R_d(z)$ can be approximated as

$$R_d(z) = R\left(\frac{2}{T_s}\frac{z-1}{z+1}\right)$$

As of now, after having selected a proper sampling time T_s, the continuous-time transfer function of the controller $R(s)$ can be translated into a discrete-time one $R_d(z)$. Similarly to the continuous-time case discussed in Chapter 3, a generic discrete-time transfer function, and in particular the one of the controller, can be written as

$$R_d(z) = \frac{b_n z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0}{z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0}$$

dividing both the numerator and the denominator by z^{-n} we further obtain

$$R_d(z) = \frac{b_n + b_{n-1}z^{-1} + \cdots + b_1 z^{1-n} + b_0 z^{-n}}{1 + a_{n-1}z^{-1} + \cdots + a_1 z^{1-n} + a_0 z^{-n}} \tag{13.1}$$

Recalling that the discrete-transfer function of the controller $R_d(z)$ describes the relationship between the discrete-time error signal $e_d(k)$ and the discrete-time control signal $u_d(k)$, i.e.

$$U_d(z) = R_d(z)\,E_d(z)$$

we can further write

$$\left(1 + a_{n-1}z^{-1} + \cdots + a_1 z^{1-n} + a_0 z^{-n}\right) U_d(z) =$$

$$= \left(b_n + b_{n-1}z^{-1} + \cdots + b_1 z^{1-n} + b_0 z^{-n}\right) E_d(z)$$

and transform back to the discrete-time domain

$$u_d(k) + a_{n-1}u_d(k-1) + \cdots + a_0 u_d(k-n) =$$

$$= b_n e_d(k) + b_{n-1}e_d(k-1) + \cdots + b_1 e_d(k-n+1) + b_0 e_d(k-n)$$

Finally, isolating the term to be computed $u_d(k)$ from the other quantities, we obtain

$$u_d(k) = -a_{n-1}u_d(k-1) - \cdots - a_1 u_d(k-n+1) - a_0 u_d(k-n) +$$

$$+ b_n e_d(k) + b_{n-1}e_d(k-1) + \cdots + b_1 e_d(k-n+1) + b_0 e_d(k-n)$$

which represents a recursive formula to compute, at discrete-time instant k, the control action $u_d(k)$ based on its previous values, i.e. $u_d(k-1), \ldots$, as well as on the values of the error $e_d(k), e_d(k-1), \ldots$.

Example 74. *[Executable code from discrete-time transfer function] Consider the following discrete-time transfer function*

$$R_d(z) = 0.7\frac{z-0.9}{z-1} = \frac{0.7-0.63z}{z-1} = \frac{0.7z^{-1}-0.63}{1-z^{-1}}$$

Then, it follows that

$$U_d(z)\left(1-z^{-1}\right) = E_d(z)\left(0.7z^{-1}-0.63\right)$$

and therefore

$$u_d(k) = u_d(k-1) + 0.7e_d(k-1) - 0.63e_d(k)$$

which clearly represents an iterative algorithm to update $u_d(k)$ based on its previous value $u_d(k-1)$ and on the present and previous values of the error, i.e. $e_d(k)$ and $e_d(k-1)$, respectively.

Back to a generic transfer function $R_d(z)$, without any lack of generality, we might implement (e.g. in C, or in any other language) the following executable code implementing the designed control functionality:

```
typedef struct
{
    float[] a;  // [a0, a1, ..., 1], length is n+1
    float[] ud; // [ud(k), ud(k-1), ...]
    float[] b;  // [b0, b1, ..., bn], length is n+1
    float[] ed; // [ed(k), ed(k-1), ...]
    unsigned int n;
} params;

void ctrl_fun(float e, params &p)
{
    // local copies
    unsigned int n = p.n;

    // left shift
    for(unsigned int k = n; k > 0; --k)
    {
        p.ed[k] = p.ed[k-1];
        p.ud[k] = p.ud[k-1];
    }

    // push front of ed(k)
    p.ed[0] = e;

    // calculation of ud(k)
    float p.ud[0] = 0.0f;
    for(unsigned int k = 0; k <= n; ++k)
    {
        p.ud[0] += b[n-k]*ed[k] - a[n-k]*ud[k];
    }
}
```

Listing 13.1: Generic implementation of a controller.

Notice that the control variable evaluated at discrete-time instant k, i.e. $u_d(k)$, will be stored in some location in memory and the corresponding value will remain unmodified until the next iteration of the controller at discrete-time instant $k+1$, thus inherently implementing the zero order holding functionality described earlier.

Example 75. *[Mass with friction] For the system in Fig. 2.7, in Chapter 8, we obtained the following controller*

$$R(s) = 4 \cdot 10^{-3} \frac{(1+100s)(50s+3)}{s^2(1+s/3.5)}$$

with $\omega_c = 0.398$ *rad/s and* $\psi_m = 82.1$ deg. *We can then try to implement its digital counterpart by selecting*

$$T_s = \frac{\pi}{50\omega_c} = 0.1579\ s$$

which produces a decrease in the phase margin of

$$\Delta\psi_m = \frac{180}{\pi}\omega_c \frac{T_s}{2} = 1.8\,\mathrm{deg}$$

In turn, the low-pass filter to remove high-frequency components of the measurement noise can be selected with a cut-off frequency of $\omega_f = 20\omega_c = 7.96$ *rad/s, corresponding to a decrease in the phase margin of*

$$\Delta\psi_m = \mathrm{atan}\left(\frac{\omega_c}{20\omega_c}\right) \approx 2.86\,\mathrm{deg}$$

We assume the overall decrease in the phase margin of $\Delta\psi_m \approx 4.66$ deg *to be reasonably low. For the discretisation of the given continuous-time controller, we obtain*

$$R_d(z) = R\left(\frac{2}{T_s}\frac{z-1}{z+1}\right) = \frac{4.353z^3 - 4.305z^2 - 4.353z + 4.305}{z^3 - 2.567z^2 + 2.134z - 0.5671}$$

Dividing both the numerator and the denominator by z^{-3} *we can rewrite* $R_d(z)$ *in terms of negative powers of* z*, i.e.*

$$R_d(z) = \frac{4.353 - 4.305z^{-1} - 4.353z^{-2} + 4.305z^{-3}}{1 - 2.567z^{-1} + 2.134z^{-2} - 0.5671z^{-3}}$$

which, finally, corresponds to the following recursive update of the control action $u_d(k)$

$$\begin{aligned} u_d(k) =& 2.567u_d(k-1) - 2.134u_d(k-2) + 0.5671u_d(k-3) + \\ & 4.353e_d(k) - 4.305e_d(k-1) - 4.353e_d(k-2) + 4.305e_d(k-3) \end{aligned} \tag{13.2}$$

Figure 13.10 reports the behaviour of the output and of the actuated force for the digital control system of Fig. 13.3 as compared to the ones produced by the original analogue control system in Fig. 7.1.

From the last example notice that, starting from a strictly proper controller $R(s)$ and applying the given discretisation method, the transfer function of the corresponding discrete-time controller $R_d(z)$ is proper, thought not strictly. As a consequence, the corresponding recursive update of the control action (13.2) presents an instantaneous dependency at (present) discrete-time instant k between the error $e_d(k)$ and the corresponding control action $u_d(k)$. This always happens whenever $b_n \neq 0$, and is surely expected whenever transfer function $R(s)$ is just proper.
No matter how long the sampling time T_s is, the evaluation of $u_d(k)$ at the

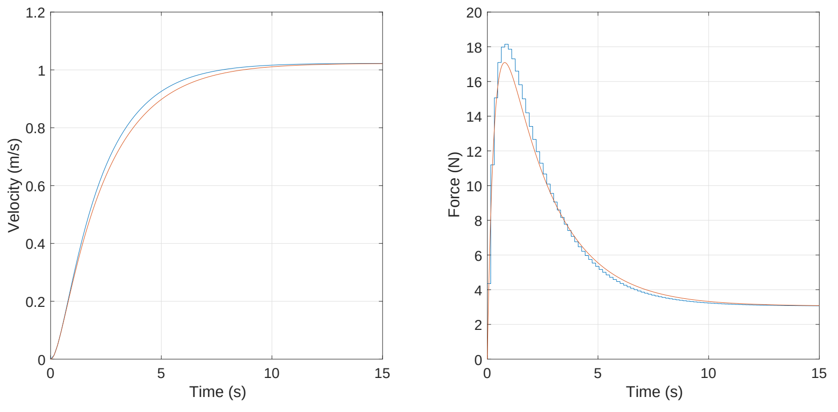


Figure 13.10: Velocity as computed by the digital (blue) and the analogue (red) control systems (left) and corresponding forces (right).

same time instant of when $e_d(k)$ is made available is clearly impossible. In a more practical scenario, the evaluation of the control action $u_d(k)$, i.e. the execution time of function `ctrl_fun` in the corresponding control code, requires T_c seconds, where, clearly, $T_c \ll T_s$.

In general, the evaluation time T_c can be regarded as a random variable as it depends on many factors, including the availability of CPU time to execute `ctrl_fun`, the presence of other tasks, interrupts, alarms, etc. Hence, in order to avoid a stochastic behaviour of the control strategy, in practical applications the value $u_d(k)$ is temporarily stored in the memory of the computing system and delivered to the actuator **at the end of the sampling period** T_s. This clearly causes an additional delay of T_s seconds between when the control variable should be ready, i.e. at discrete-time instant k, and when it is actually delivered to the actuator, i.e. at discrete-time instant $k+1$. The situation is sketched in Fig. 13.11.

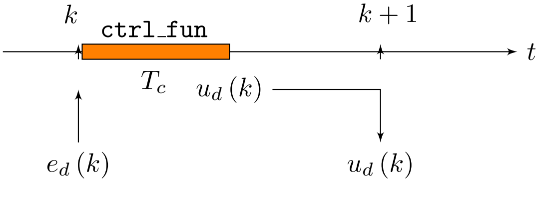


Figure 13.11: Schematics of the execution of the control code `ctrl_fun`.

The overall block diagram for the digital implementation of the controller is reported in Fig. 13.12. Notice that the transfer function of the controller $R_d(z)$ is already including a delay of one sampling interval (i.e. z^{-1}) consistently with the discussion above.

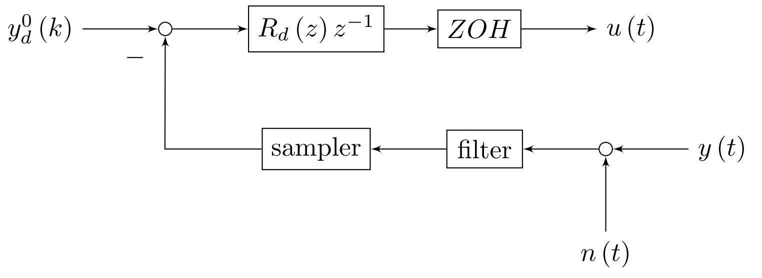


Figure 13.12: Complete block diagram for the implementation of a digital control law.

Part IV

Technologies for digital feedback control systems

Chapter 14

Hardware technologies for feedback control systems

This Chapter is focused on hardware technologies for the implementation of a digital control system and in particular the one in Fig. 13.12. We will address all the technologies behind the D/A and the A/D conversions.

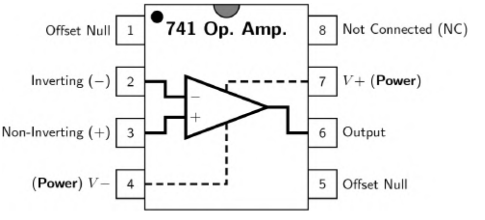
14.1 The operational amplifier

The building block of all the hardware needed to implement control-related functionalities is the **operational amplifier** (or op-amp).
An operational amplifier is a power-supplied differential voltage amplifier. When supplied with voltages $V^+ > V_-$, the output $v_{out}(t)$ is provided as

$$v_{out}(t) = A\left(v^+(t) - v^-(t)\right) \tag{14.1}$$

where A is the gain of the amplifier (typically a high value in the range $10^4 - 10^5$), while $v^+(t) - v^-(t)$ is the differential input voltage.

(a) Operational Amplifier (ST Microelectronics)



(b) Corresponding pinout

Figure 14.1: IC Operational Amplifier LM741 and corresponding pinout.

An example of operational amplifier is reported in Fig. 14.1 together with the corresponding pinout. The symbolic representation of an op-amp is reported in Fig. 14.2.

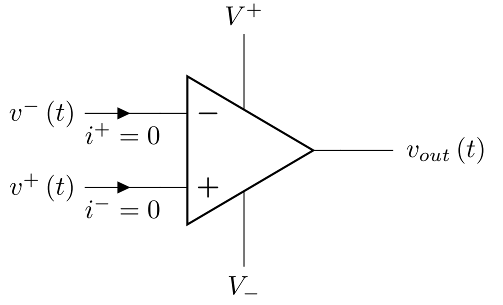


Figure 14.2: An operational amplifier (op-amp).

Due to its manufacturing characteristics, the input ports absorb no current. Moreover, for a non-zero value of the differential input voltage $v^+(t) - v^-(t)$, the output voltage $v_{out}(t)$ easily saturates to either the positive voltage supply V^+, or to the negative one V_-, depending on the sign of $v^+(t) - v^-(t)$. In first approximation

$$v_{out}(t) = A\left(v^+(t) - v^-(t)\right) \approx \begin{cases} V^+ & v^+(t) > v^-(t) \\ V_- & \text{otherwise} \end{cases}$$

In the arrangement of Fig. 14.2, the operational amplifier acts as a comparator between the two input voltages $v^+(t)$ and $v^-(t)$.

14.1.1 Operational amplifiers with negative feedback

Inverting amplifier

Consider the operational amplifier in the arrangement shown in Fig. 14.3 where it is easy to realise the negative feedback arrangement as the output voltage $v_{out}(t)$ is fed back to the negative input port. Since, as anticipated, the input ports draw no current, the current flowing in R_1 will also flow in R_2, i.e.

$$i_1(t) = \frac{v_{in}(t) - v^-(t)}{R_1} = i_2(t) = \frac{v^-(t) - v_{out}(t)}{R_2} \tag{14.2}$$

and therefore

$$(R_1 + R_2)\, v^-(t) = R_2 v_{in}(t) + R_1 v_{out}(t)$$

From (14.1) and since $v^+(t) = 0$, and therefore $v^-(t) = -v_{out}(t)/A$, we obtain

$$-\frac{R_1 + R_2}{A} v_{out}(t) = R_2 v_{in}(t) + R_1 v_{out}(t)$$

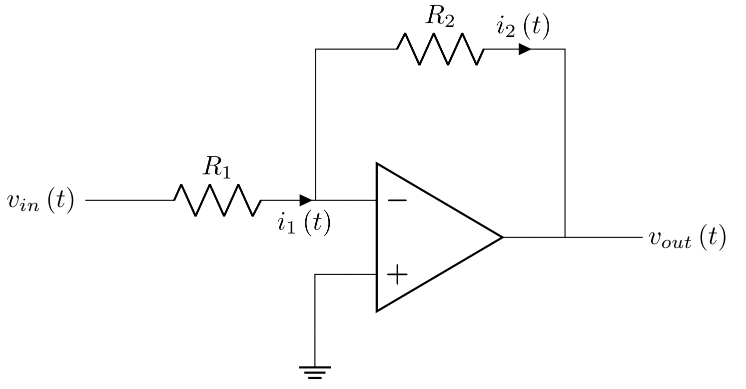


Figure 14.3: An operational amplifier with negative feedback (inverting amplifier).

and finally

$$v_{out}(t) = -R_2 \left(R_1 + \frac{R_1 + R_2}{A} \right)^{-1} v_{in}(t)$$

Since the operational amplifier is characterised by very high gains, we have

$$v_{out}(t) \approx -R_2 \lim_{A \to +\infty} \left(R_1 + \frac{R_1 + R_2}{A} \right)^{-1} v_{in}(t) = -\frac{R_2}{R_1} v_{in}(t)$$

Because of the last relationship between the input voltage $v_{in}(t)$ and the output voltage $v_{out}(t)$, the arrangement of Fig. 14.3 is called **inverting amplifier**. Finally, from (14.2), we can also verify that

$$v^-(t) \approx 0 = v^+(t)$$

The last approximation is also called **virtual ground** as the voltage of the negative port $v^-(t)$ is (almost) equal to the one of the positive $v^+(t)$, which is the ground voltage, at least in this setting. Hence, still referring to the inverting amplifier in Fig. 14.3, the negative port has the same voltage of the ground, though not being directly connected to it.

Non-inverting amplifier

The arrangement of Fig. 14.4 is called **non-inverting amplifier**. The relationship between the input voltage $v_{in}(t) = v^+(t)$ and the output voltage $v_{out}(t)$ can be computed resorting to the virtual ground principle just introduced

$$v_{out}(t) = \left(1 + \frac{R_2}{R_1} \right) v_{in}(t)$$

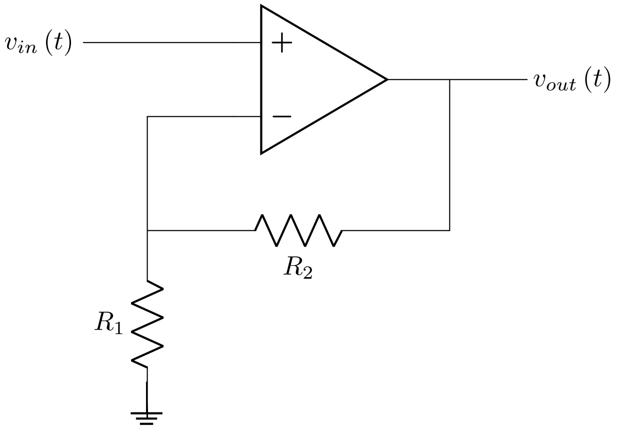


Figure 14.4: Non-inverting amplifier.

Moreover, letting $R_2 = 0$ and $R_1 \to +\infty$, the relationship becomes

$$v_{out}(t) = v_{in}(t)$$

which is called **voltage buffer** (see Fig. 14.5) and is primarily adopted to galvanically isolate the input from the output.

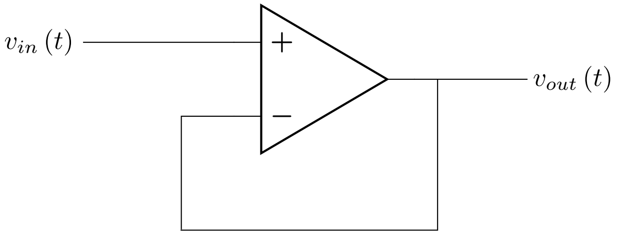


Figure 14.5: Voltage buffer.

Summing amplifier

The arrangement of the operational amplifier in Fig. 14.6 is called **summing amplifier**. The current flowing in each leg $i_i(t)$ connected to the negative port depends on the corresponding input voltage $v_i(t)$ and on the corresponding resistance R_i. The sum of these currents then flows in resistance R_f. The relationship between the output voltage $v_{out}(t)$ and the inputs voltages $v_1(t), \ldots, v_n(t)$ is as follows

$$v_{out}(t) = -R_f \sum_{i=1}^{N} \frac{v_i(t)}{R_i}$$

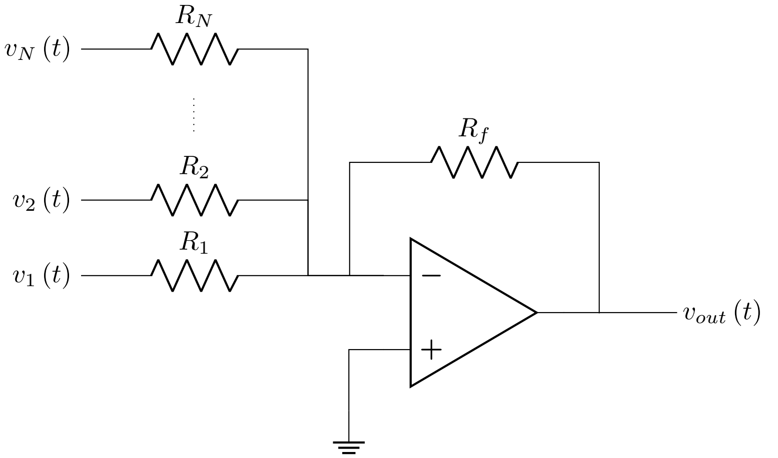


Figure 14.6: Summing amplifier.

14.2 Operational amplifiers for filtering

As of now, we have only introduced algebraic relationships between some input voltages and the output voltage $v_{out}(t)$. In case one or more resistors are substituted with dynamic elements, such as capacitors, the relationship between the input(s) and the output becomes dynamic.

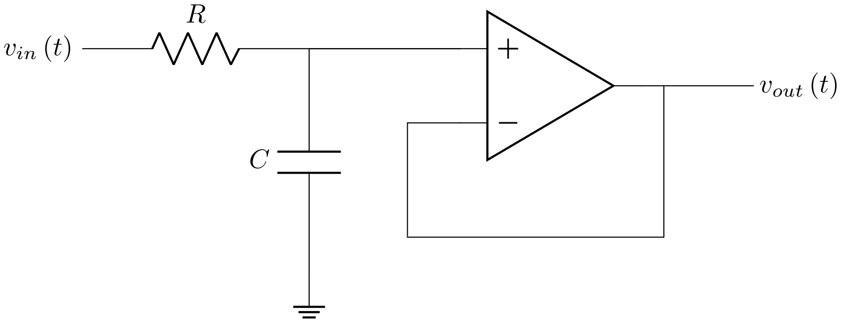


Figure 14.7: Low-pass filter.

Consider the arrangement of Fig. 14.7. The relationship between the input and the output voltage, in terms of their Laplace's transforms, is the following one

$$V_{out}(s) = \frac{1}{1 + sRC} V_{in}(s)$$

which clearly represents a low-pass filtering operation on signal $v_{in}(t)$. As seen in Chapter 13, the low-pass filtering functionality is extremely important to remove high-frequency components of the measured signal before being sampled.

Example 76. *Consider the low-pass filter of Fig. 14.7 to be used to remove high-frequency components of the measurement noise in the last example of Chapter 13 that had to be designed with a cut-off frequency of* $\omega_f = 7.96$ *rad/s.*
It should result immediate to set $RC = 1/\omega_f$*, i.e. for example* $R = 1$ $k\Omega$*, and* $C = 125.63$ μF*.*

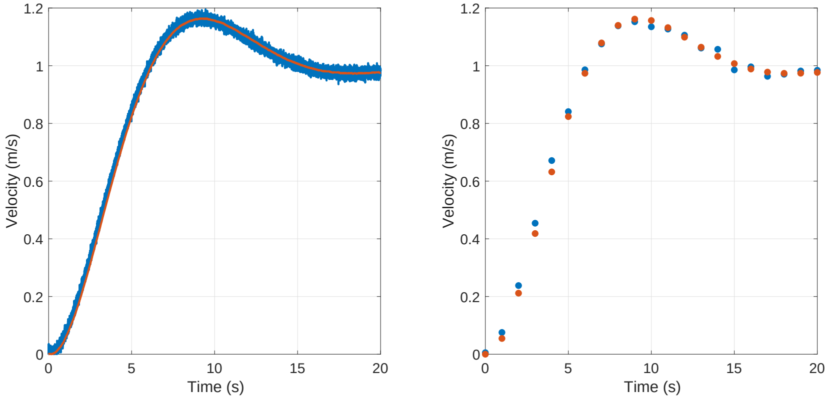


Figure 14.8: Acquired (blue) and filtered velocities (red) on the left, and corresponding samples (right).

Figure 14.8 reports the behaviour of the measured velocity with and without the low-pass filter as well as the corresponding samples.

14.3 Operational amplifiers for D/A conversion

In Chapter 13 we introduced the D/A converter as one of the element interfacing the digital controller with the physical world. In particular, we described the D/A in terms of its behaviour without further details regarding its physical implementation. In this Section, we further investigate the D/A, but in terms of constructing technologies. The role of the D/A is twofold: convert a series of bits into an analogue signal, and to maintain such a signal constant along the sampling interval T_s.
The first problem is addressed using operational amplifiers. Assume that the control signal $u_d(k)$ is encoded in N bits. The arrangement of the summing amplifier of Fig. 14.6 is particularly interesting for this problem. Assume that the resistances $R_1, R_2, \ldots, R_N$ are arranged in terms of powers of two, i.e. $R_i = 2^{1-i} R_f, i = 1, \ldots, N$. The relationship between the input voltages

$v_i(t), i = 1, \ldots, N$ and the output voltage $v_{out}(t)$ is given in the following:

$$\begin{aligned} v_{out}(t) &= -R_f \sum_{i=1}^{N} \frac{v_i(t)}{2^{1-i} R_f} = -\sum_{i=1}^{N} \frac{v_i(t)}{2^{1-i}} = \\ &= -\left(v_1(t) + 2v_1(t) + 4v_2(t) + \cdots + v_N(t)\, 2^{N-1}\right) \end{aligned} \tag{14.3}$$

It follows that, if the input voltages $v_i(t), i = 1, \ldots, N$ are set to represent the string of bits encoding $u_d(k)$, the output voltage $v_{out}(t)$ will represent the corresponding analogue value, apart from the sign.

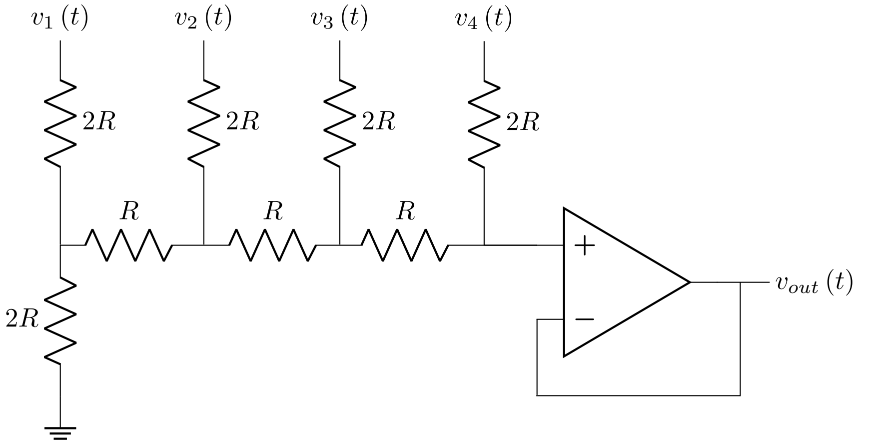


Figure 14.9: $R - 2R$ 4 bits D/A converter.

The methodology just introduced to convert a string of bits into an analogue voltage can work, provided that a set of resistors properly scaled with the powers of two can be found. Notice that the ratio between the smallest and the largest resistances is 2^{N-1}. For a D/A converter with $N = 12$ bits, this value would be $2^{11} = 2048$, meaning that resistances span more than three orders of magnitude.

A better and more practical solution is to rely on the so-called $R - 2R$ ladder circuit which is reported in Fig. 14.9. For a generic number of bits N, the relationship between the input voltages $v_1(t), v_2(t), \ldots, v_N(t)$ and the output voltage $v_{out}(t)$ is the following one

$$v_{out}(t) = 2^{-N} \sum_{i=1}^{N} 2^{i-1} v_i(t)$$

Hence, once again, if the input voltages $v_i(t), i = 1, \ldots, N$ are set to represent the string of bits encoding $u_d(k)$, the output voltage $v_{out}(t)$ will represent the corresponding analogue value. Notice that, differently from the summing amplified, the $R-2R$ ladder is using just two types of resistors, one valued R, the other $2R$.

Finally, some remarks can be given regarding the zero order holding (ZOH) functionality that is typically present in D/A converters. To this end, notice that at any discrete instant k, the bits representing the actuation $u_d(k)$ are stored in some memory. Therefore, **the memory itself will already play the role of the holder** by maintaining the value stored until the next cycle when it will be overwritten by the updated control command $u_d(k)$.

14.4 Operational amplifiers for A/D conversion

As opposed to the D/A conversion, the role of the A/D converter is threefold: filtering high-frequency noise typically characterising the measured variables, sampling the filtered value every T_s seconds, and the actual digital conversion of the sample from analogue to digital. The first role has been already described within the discussion on operational amplifiers arranged to act as filters (refer again to Fig. 14.7). In the next we focus on the actual A/D conversion which deserves more detailed explanations.

As the name suggests, A/D conversion is the inverse of the D/A operation and consists in producing an analogue voltage corresponding to a certain string of bits. We have already seen that, within a feedback controlled system, the relationship between the reference signal $y^0(k)$ and the control variable $u(k)$ is expressed by means of the **control sensitivity**, i.e.

$$U(z) = \frac{R(z)}{1 + R(z)G(z)} Y^0(z)$$

where $G(z)$ is the system to be controlled and $R(z)$ is the transfer function of the regulator. We have also seen that, for low frequencies, the control sensitivity approximates the inverse of $G(z)$, i.e.

$$U(z) = \frac{R(z)}{1 + R(z)G(z)} Y^0(z) \approx \frac{1}{G(z)} Y^0(z)$$

Therefore, apart from fast transients, the control variable $u(k)$ is related to $y^0(k)$ through the inverse of the system under control, i.e. $G(z)^{-1}$. The first idea on how to obtain the A/D functionality is then to feedback control a D/A converter.

The functional diagram of a generic A/D converter is reported in Fig. 14.10. The constant voltage v_{in} to be converted (which plays the role of the constant reference y^0) is first compared against the output of a D/A converter and, depending on the outcome of the comparison, a control law iteratively adjusts the

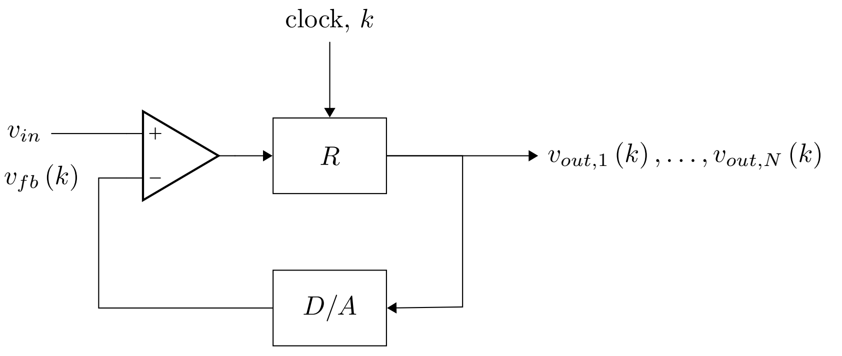


Figure 14.10: Successive approximation register (SAR) A/D converter.

input of the D/A converter until convergence.

More specifically, the conversion operates in successive iterations as described in the following. At the beginning, i.e. for $k = 0$, all bits are initially set to low, i.e. $v_{out,1}(0) = 0, \dots, v_{out,N}(0) = 0$, then from the most to the least significant bits (MSB and LSB, respectively), at each iteration $k = 1, \dots, N$

1. the **control logics** contained in R sets the k-th bit to high, i.e.

$$v_{out,N-k+1}(k) = 1$$

2. the D/A converter performs the conversion producing $v_{fb}(k)$;

3. finally, if $v_{fb}(k) < v_{in}$, the control logics R confirms the bit as high, $v_{out,N-k+1}(k+1) = 1$, otherwise (i.e. if $v_{fb}(k) \geq v_{in}$) it sets the bit to low, i.e. $v_{out,N-k+1}(k+1) = 0$.

The strategy just described to implement the A/D conversion is called **successive approximation register** (SAR). The reader should not be surprised in observing that the control logics contained in a SAR A/D converter actually constitutes a discrete-time dynamical system. In fact, letting $\boldsymbol{v}_{out}(k) = \begin{bmatrix} v_{out,1}(k) & \dots & v_{out,N}(k) \end{bmatrix}^T$ the state of the control logics, the algorithm just introduced can be rewritten as follows

$$v_{out,i}(k+1) = \begin{cases} v_{out,i}(k) & i < k \\ v_{fb}(k) < v_{in} & i = k \\ 0 & \text{otherwise} \end{cases}$$

Example 77. *[A/D conversion based on SAR] Consider a $N = 8$ bits SAR A/D converter fed with an input voltage of $v_{in} = 5.3$ V and a maximum voltage of $V^{\max} = 8$ V.*

Within the first iteration, $k = 1$, the MSB is set to high, that is 10000000. The corresponding analogue voltage

$$v_{fb}(1) = \frac{2^7}{2^8} 8 = 4V$$

is then produced by the D/A. Since $v_{fb}(1) < v_{in}$, the bit will be confirmed as high in the next iteration.

Then, in the second iteration, the controller R sets the second bit to high, i.e. 1100000 and voltage

$$v_{fb}(2) = \frac{2^7 + 2^6}{2^8} 8 = 6V$$

which is higher than the reference voltage v_{in}, hence the bit will be changed to low.

The iterations continue until the last bit is set either to low or to high. After $k = N = 8$ iterations, the resulting value for the A/D conversion is 10101001 which corresponds to

$$v_{fb}(8) = \frac{2^7 + 2^5 + 2^3 + 2^0}{2^8} 8 = 5.2812V \approx 5.3V$$

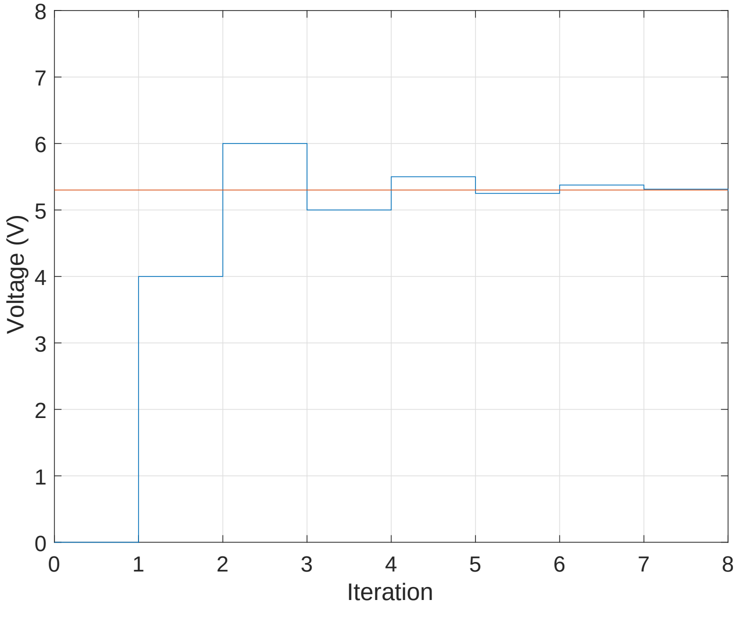


Figure 14.11: Intermediate values of $v_{fb}(k)$ (blue) produced by the SAR A/D converter and corresponding reference value v_{in} (red).

Figure 14.11 reports the evolution of the voltage $v_{out}(k)$ produced by the D/A for every iteration k as compared to the (constant) reference voltage v_{in}.

From the last example, it should be clear that the conversion of the analogue voltage $v_{in}(t)$ into an equivalent string of N bits is an iterative process which is performed in N steps and the overall process has to complete in way less than the sampling interval T_s. Therefore, letting T_c the time required by the D/A converter to produce the output voltage $v_{out}(N)$, the overall process requires $NT_c \ll T_s$ seconds. Within this time interval, the reference voltage $v_{in} \in [-V^{\max}, V^{\max}]$ has to be maintained as constant as possible, or at least within the quantisation interval $2V^{\max}/2^N$. Since a low-pass filter with cut-off frequency of ω_f is always available to filter out high-frequency components of the measurement noise, we can use ω_f as a proxy for the variability of $v_{in}(t)$. Then assuming

$$|\Delta v_{in}(t)| = NT_c V^{\max} \omega_f$$

to be the maximum variability of $v_{in}(t)$ within the conversion interval NT_c, in order to correctly operate the A/D conversion we need

$$|\Delta v_{in}(t)| = NT_c V^{\max} \omega_f < \frac{2V^{\max}}{2^N}$$

and therefore

$$NT_c < \frac{1}{2^{N-1}\omega_f}$$

For a given A/D converter, hence for given T_c and N, if the aforementioned inequality is not satisfied, the simple holding circuit of Fig. 14.12 is placed before the A/D converter.

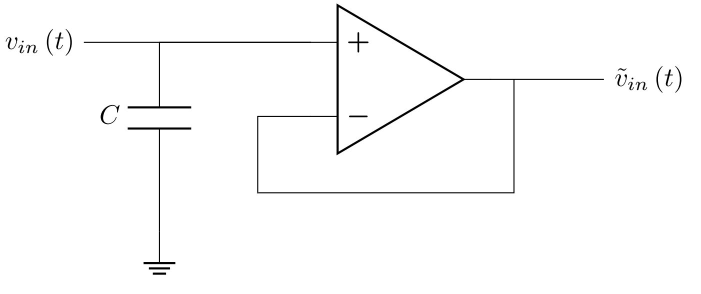


Figure 14.12: Holding circuit.

The holding circuit of Fig. 14.12 basically stabilises its output voltage $\tilde{v}_{in}(t)$ by removing possible fluctuations of $v_{in}(t)$ that are higher than the quantisation interval $2V^{\max}/2^N$.

Finally, the sampling operation is simply obtained using a switch (typically implemented using transistors) before the holding circuit of Fig. 14.12 which commutes and closes every T_s seconds for the time necessary to charge/discharge the capacitor of the holding circuit.

For particularly demanding application, the **flash A/D converter** constitutes a valid alternative to the SAR converter. A flash converter operates in a one-shot fashion producing the conversion in T_c seconds, independently from the number N of bits. The corresponding circuit is shown in Fig. 14.13. The left leg of the circuit represents a voltage divider which sets the $2^N - 1$ negative ports of the corresponding operational amplifiers (ordered from bottom to top) to the following values:

$$v_i^- (t) = \frac{R/2 + (i-1) R}{(2^N - 1) R} V^{\max} = \frac{i - 1/2}{2^N - 1} V^{\max}, i = 1, \ldots, 2^N - 1$$

It is worth noticing that these value are monotonically increasing with i, i.e.

$$v_{i+1}^- (t) > v_i^- (t), \forall i = 1, \ldots, N - 1$$

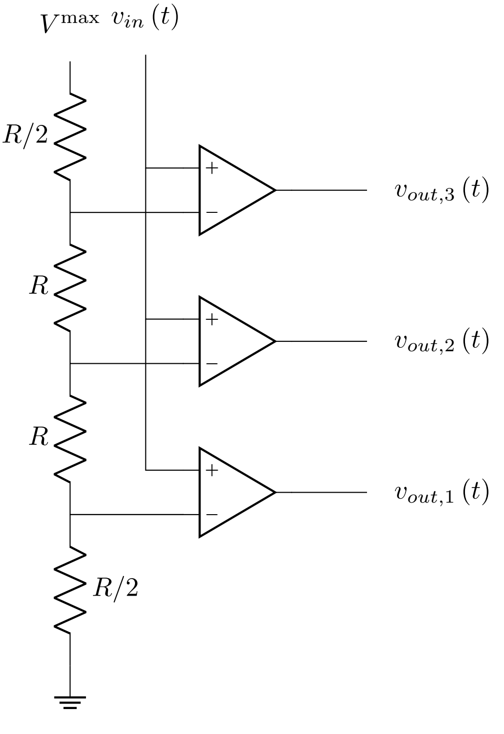


Figure 14.13: A flash D/A converter with $N = 2$ bits.

The operational amplifiers are here adopted as comparators, therefore the logical values of their outputs will be set to

$$v_{out,i} (t) = \begin{cases} V^{\max} & v_{in} (t) > \frac{i - 1/2}{2^N - 1} V^{\max} \\ 0 & \text{otherwise} \end{cases}, i = 1, \ldots, 2^N - 1$$

forming the so-called **thermometer encoding** of the input voltage $v_{in}(t)$. The reason can be simply explained by noticing that if $v_{out,i}(t)$ is set to a high value, then all $v_{out,j}(t), j \leq i$ will be set to high, as well.
The actual digital encoding of voltage $v_{in}(t)$ is finally performed by a subsequent circuit that simply counts the output voltages that were set to high values and converts this count in standard binary code (see Tab. 14.1).

$v_{out,3}$	$v_{out,2}$	$v_{out,1}$	count	binary
0	0	0	0	00
0	0	1	1	01
0	1	1	2	10
1	1	1	3	11

Table 14.1: Conversion from thermometer to standard binary encoding for $N = 2$ bits.

Example 78. *Consider a $N = 8$ bits flash A/D converter fed with the constant input voltage of $v_{in} = 5.3\ V$ of the previous example, and a maximum voltage of $V^{\max} = 8\ V$. The negative ports of the $2^N - 1 = 255$ operational amplifiers are fed with the following reference voltages*

$$v_1^-(t) = \frac{1 - 1/2}{2^8 - 1}8 = 0.0157\ V, v_2^-(t) = \frac{3/2}{2^8 - 1}8 = 0.0471\ V, \ldots$$

$$v_{169}^-(t) = \frac{169 - 1/2}{2^8 - 1}8 = 5.2863\ V, v_{170}^-(t) = \frac{170 - 1/2}{2^8 - 1}8 = 5.3176\ V,$$

$$\ldots, v_{255}^-(t) = 7.9843\ V$$

Therefore the binary encoding of the analogue voltage $v_{in} = 5.3\ V$ will be 10101001, i.e. binary(169).

Notice that, since the $2^N - 1$ operational amplifiers composing the flash converter operates in parallel, the A/D conversion time is just equal to the time T_c required by the comparison making this converter significantly faster than the SAR, especially for high numbers of bits N.

14.5 Summary

In this Chapter, we have seen how most of the technologies required for the implementation of the digital control system in Fig. 13.3 are physically implemented in hardware by means of operational amplifiers. Figure 14.14 reports an extremely simplified version of a digital controller comprising the A/D and the D/A functionalities. As also discussed in Chapter 13, a low-pass filter is also typically deployed to remove high-frequency noise in the measured signal. Moreover, a non-inverting amplifier is placed before the sampler in order to scale the filtered signal to fit the fully range of the A/D converter.

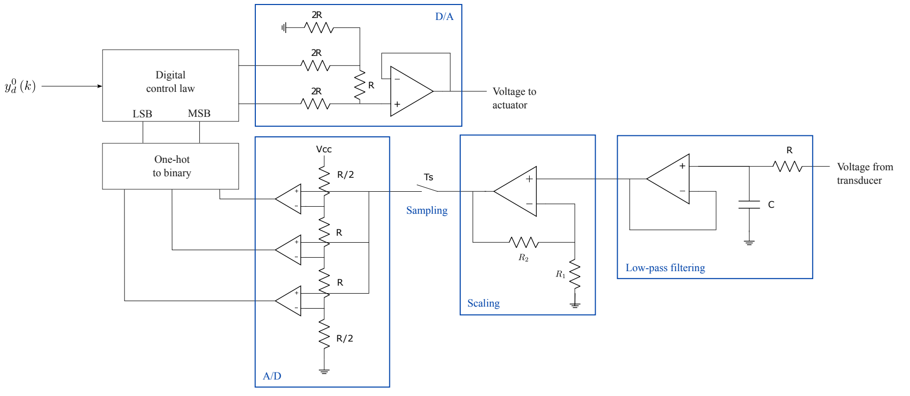


Figure 14.14: Schematic representation of the pipeline for a $N = 2$ bits controller.

The software functionalities required for the implementation of the digital control law, which can be derived from the discrete-time controller transfer function $R_d(z)$ as seen in Chapter 13, will be further discussed in the next Chapter.

Chapter 15

Computing technologies for feedback control systems

The previous Chapters of this book were essentially focused on the methods and the corresponding technologies for the implementation of a feedback strategy to let a single output variable $y(t)$ follow an independent reference $y^0(t)$.
Consider an automated warehouse as the one reported in Fig. 15.1. In order to retrieve the correct box, the picker has to first transfer towards the correct shelf, then move vertically to the correct ledge, then extract the picking clamps, close and retract them, and move to the drop-off station.

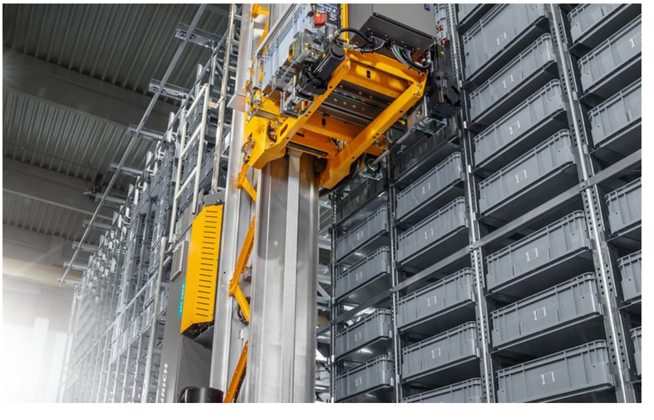


Figure 15.1: An automated warehouse (Jungheinrich).

As we might argue, to govern the entire system a single control loop is not sufficient. Surely, different controllers should be employed: one to govern the motion of the picker along the shelves, one to govern its motion in the vertical direction, another one to extract and retract the picking clamps, etc. While

the corresponding control loops can be probably designed or tuned separately from each other, during operations they have to be **orchestrated** to achieve the desired goal. For example, the clamping device should be extracted **only when** the vertical position of the picker has reached its steady-state value corresponding to the desired ledge. The clamps can be activated **only when** the device carrying them has been fully extracted, etc. From this simple example, we should notice that a relatively simple automated control system actually requires both the regulation of continuous-time physical quantities (e.g. the horizontal and vertical positions of the picker) as well as a proper **action sequencing** strategy. Similar examples can be mentioned in other fields of engineering. For example, an automated underground train has to open its doors **only when** the train has reached the next station and its velocity is reasonably null. In turn, it is allowed to close the door **only after** a certain time, and **only if** no more passengers are occupying the doorways.
In the reminder of this Chapter, the software technologies for the development of such orchestration functionalities will be discussed.

15.1 Programmable Logic Controllers

A Programmable Logic Controller (PLC, in short, see also Fig. 15.2) is a programmable rugged device offering extensive input/output (I/O) functionalities to handle the connections from sensors and towards actuators.

Figure 15.2: A Programmable Logic Controller (Allen-Bradley).

Every PLC program is executed repeatedly, every T_s seconds (each program has its own T_s). Differently from programs meant for the execution on standard computers, a PLC program is constituted by the following steps:

1. the levels of physical inputs is copied to an area of memory accessible to the CPU (I/O Image Table);
2. the user program is run, and the I/O Image Table is updated with the new values of the corresponding outputs (refer again to the schematics of Fig. 13.11);

3. operating system services are executed (for scheduling purposes, mainly);
4. the outputs are copied from the I/O Image Table to physical outputs.

Under the IEC 61131-3:2013 standard, PLCs can be programmed using standardised programming languages:

- Function Block Diagram (FBD);
- Ladder Diagram (LD);
- Sequential Function Chart (SFC);
- Structured Text (ST or STX).

FBD, LD and SFC are graphical programming languages, while ST is a textual language similar to PASCAL. Additional programming languages such as specially adapted dialects of BASIC and C might be also available, in some cases. As we have already addressed in Chapter 13, a discrete-time transfer function $R_d(z)$ can be easily translated into an algorithm to be executed. Then, assuming Structured Text to be the programming language for its implementation, the corresponding code can be easily written. As a matter of fact, the majority of commercial PLCs offer a wide range of pre-implemented controllers, obviously including those belonging to the PID family.

15.2 Sequential Function Charts

As for the typical orchestration functionalities required in almost all automatic control applications, graphical programming languages are the best candidates. **Sequential Function Chart** (SFC) is a graphical programming language that can be used to control processes that can be split into different execution steps. The main (graphical) components of a SFC program are:

- **steps** with associated actions;
- **transitions** together with the associated logical conditions;
- directed links between steps and transitions, or between transitions and steps.

Steps indicate operational condition of both the controller and the system under control and are graphically represented by squared boxes. A step can be either active or inactive (when active the squared box contains a marker). The marker is just a graphical representation of a Boolean variable indicating whether the step is active or not. Actions are only executed for active steps. Activation of a step happens for one of the two following reasons:

- it is an initial step (represented by a double-bordered squared box) as specified by the programmer;

- it has been activated during a previous PLC cycle and not deactivated since.

Steps are activated when all steps above it are simultaneously active and the connecting **transition** is superable (i.e. its associated logical condition is true). A transition is represented with a horizontal bar. When a transition is passed, all steps before the transition are deactivated (corresponding markers are removed) at once and all steps after the transition are activated at once (corresponding markers are added).

Actions associated to steps can be of several types (named **qualifiers**), the most relevant ones being Normal or Non-stored (N), and Pulse (P). The former indicates that the corresponding action has to be continuously executed as long as the step is active. The latter, in turn, indicates that the corresponding action has to be executed just once, right after the activation of the corresponding step.

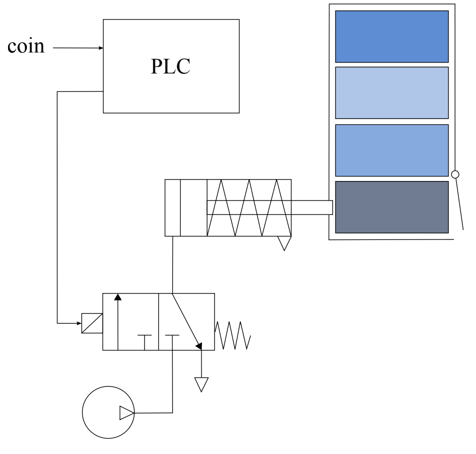


Figure 15.3: A pneumatic vending machine controlled by a PLC.

As an example, consider the pneumatic vending machine of Fig. 15.3. Upon the insertion of a coin, which is detected by means of a proper sensor acquired by the PLC, the solenoid has to be actuated, thus commuting the state of the 3-ways 2-positions normally closed valve and letting the air flow from the compressor to the normally closed single-effect piston. The piston elongates, pushing the purchased item out of the stack while preventing the next one to fall. After two seconds, the solenoid can be deactivated, the piston retracts thanks to the return spring and the system returns to its initial condition.

The system can be governed by a PLC running the SFC program sketched in Fig. 15.3. It is composed of two steps $X0$ (initial step) and $X1$ and two transitions. At the beginning the initial step $X0$ is active, and the solenoid is not energized ($solenoid = 0$). Whenever a coin is inserted ($coin = 1$) the transition after $X0$ becomes superable. Correspondingly, $X0$ is deactivated and $S1$ activated. With the activation of step $X1$, the solenoid is energized ($solenoid = 1$), thus actuating the piston. After 2 seconds, the transition after $S1$ is superable ($X.t >= t\#2s$ indicates that step $X1$ has been active for at least 2 seconds), hence $X1$ is deactivated, $X0$ activated and the cycle restarts.

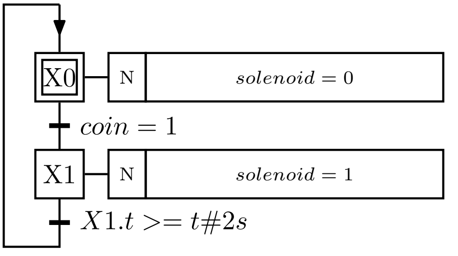


Figure 15.4: SFC code to govern the pneumatic vending machine.

Defining the state of the SFC program as $x(k) \in \{0, 1\} = \mathbb{B}$, corresponding to $X0$ active and $X1$ inactive and vice-versa respectively, its inputs as $y_1(k) = coin \in \mathbb{B}$ and $y_2(k) = X1.t \in \mathbb{R}$, and the corresponding output as $u(k) = solenoid \in \mathbb{B}$, it should not be surprising that the SFC program in Fig. 18.9 can be alternatively represented in terms of the equations

$$\begin{aligned} x(k+1) &= f(x(k), \boldsymbol{y}(k)) \\ u(k) &= x(k) \end{aligned} \tag{15.1}$$

where

$$f(x(k), \boldsymbol{y}(k)) = \begin{cases} 0 & x(k) = 1 \text{ and } y_2(k) \geq 2 \\ 1 & x(k) = 0 \text{ and } y_1(k) = 1 \\ x(k) & \text{otherwise} \end{cases}$$

which constitute a discrete-time system. Notice that both the state $x(k)$, the input $\boldsymbol{y}(k)$ and the output $u(k)$, are constituted by discrete quantities, only. In this cases, the corresponding dynamical systems are rather classified as *Discrete Event Systems* (DES), for which a separate literature exists.

15.2.1 Program flow

SFC control programs are usually more complicated than the one exemplified in Fig. 15.4. As in all programming languages, dedicated structures to modify the flow of the program are available.

If-then-else-endif

As it is often necessary to take different actions depending on some condition of the system under control, conditional branches in the form of if-then-else structures are possible in SFC programs. In the example of Fig. 15.5, whenever state $X0$ is active, either the transition with associated condition A or the other with condition $\neg A$ are superable. In both the cases $X0$ will be deactivated. In the former case, the SFC program will proceed activating $X1$, in the latter, $X3$ will be activated, instead.

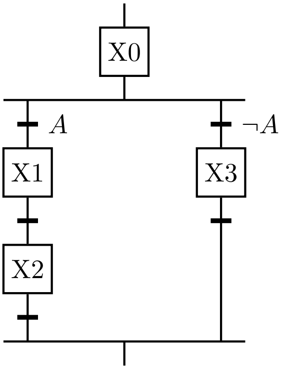


Figure 15.5: SFC excerpt for the if-then-else-endif structure.

Parallelism and rendezvous

For complex automation systems, it is usually necessary to perform two or more actions concurrently. Differently from the two previous examples in Figs. 15.4 and 15.5, in this case two or more steps can be simultaneously active. This is not in contrast with the execution paradigm for SFC program.

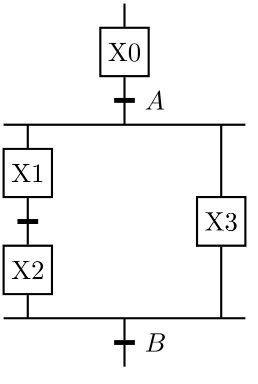


Figure 15.6: SFC excerpt for parallelism and rendezvous.

The SFC program in Fig. 15.6 contains two branches which are simultaneously (and independently) executed upon the deactivation of $X0$. The leftmost will first activate $X1$ and then $X2$, the rightmost just $X3$. Whenever the transition with associated logical condition B is superable, $X2$ and $X3$ are deactivated and the program returns to follow a single stream.

Synchronisation

Focus again on the parallel execution of two branches. As we have seen in the example of Fig. 15.6, the two (or more) parallel streams can proceed independently from each other. Assume now that one wants to force one of the two sequences to wait for the other before proceeding.

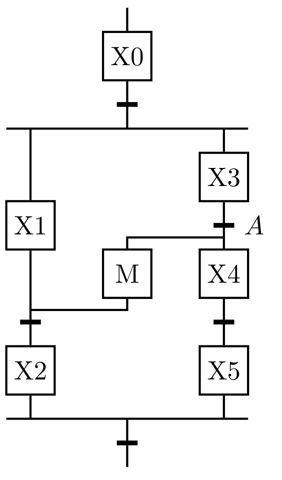


Figure 15.7: SFC excerpt for the synchronisation of two concurrent streams.

In the example of Fig. 15.7, the two parallel branches are simultaneously activated upon the deactivation of $X0$ and starts proceeding in parallel with both $X1$ and $X3$ simultaneously active. The transition between $X1$ and $X2$ is not superable until step M is active which, in turns, happens when the transition with associated condition A is superable. It follows that the activation of step $X2$ has to wait for the deactivation of $X3$.

Example 79. *[Synchronisation of two cyclic tasks] Consider two cyclic tasks (A and B) composed by 3 and 2 operations (A.1, A.2, A.3, and B.1, B.2), respectively, to be executed one after the other. And assume that B.1 cannot start before A.1 has finished, and A.3 cannot start before B.1 has finished.*

Instead of directly sketching the overall SFC program in just one shot, we can proceed gradually by first implementing the code to govern the independent execution of the two tasks. The corresponding SFC program is reported in Fig. 15.8. Notice that a step per each operation has been introduced, as well as intermediate steps to explicitly indicate waiting (idle) states. Moreover, actions and corresponding qualifiers associated to $A1, A2, A3$ *and* $B1, B2$ *are not reported.*

In order to realise the required synchronisation, we add two more steps $M1$ *and* $M2$*. Step* $M1$ *is activated upon the deactivation of* $A1$ *to indicate that* $B1$ *can start. Similarly,*

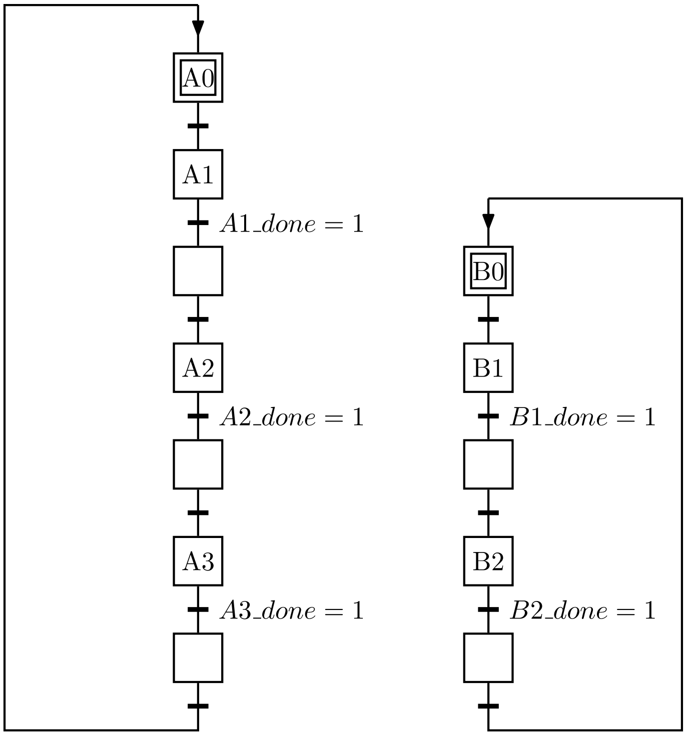


Figure 15.8: SFC code to execute the two cyclic tasks (without synchronisation).

step $M2$ will be activated upon the deactivation of $B1$ indicating that $A3$ can start. The overall SFC program is reported in Fig. 15.9.

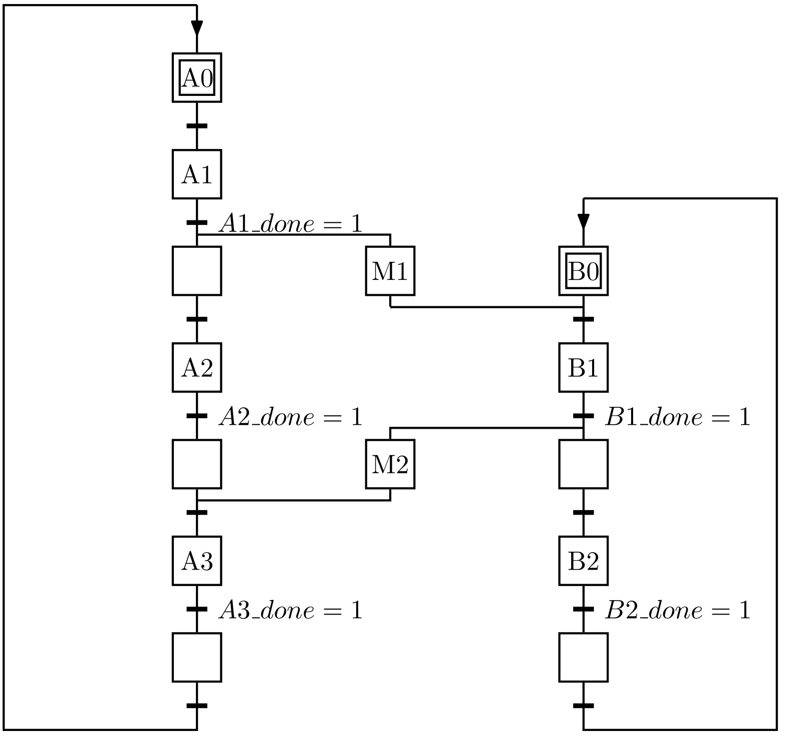


Figure 15.9: SFC code to execute the two cyclic tasks together with the prescribed synchronisation.

Part V

Mechatronic applications of control

Chapter 16

Motion control with electro-hydraulic actuators

The problem of controlling the motion of a mechanical system actuated by an electro-hydraulic actuator entails a series of methodological and technological challenges. This Chapter investigates the motion control problem for mechanical system with electro-hydraulic actuators. Mechanical systems governed by electro-hydraulic actuators are very common in different fields of engineering, especially in aeronautics. For example, control surface is indeed governed by electro-hydraulic actuators (see Fig. 16.1 as an example). Even when aircraft

Figure 16.1: Fully extended flap of a Boeing 777 during landing (Creative Commons).

are manually actuated by pilots, modern flight controllers actually adopt fly-by-wire technologies, which are essentially feedback control systems following references provided by the pilots.

16.1 Components for motion control

In the following, we will make reference to a very generic mechanism composed by an electrical servo-valve, together with its driver, a hydraulic cylinder, and a generic mechanical load. The position of the moving load is assumed to be measured by a proper position sensor.

16.1.1 Modelling the electro-hydraulic dynamics

Without any lack of generality, we will make reference to a 5-ways proportional spool servo-actuated valve. With reference to the model in Fig. 16.2, define M the mass of the mechanical load, also comprising the one of the piston and of the rod, $x(t)$ its absolute position, $q_1(t)$ the flow towards the left chamber of the cylinder, $p_1(t)$ the pressure in the left chamber, $q_2(t)$ the flow from the right chamber away from the cylinder, $p_2(t)$ the pressure in the right chamber. Further define the two constants p_d and p_e being the delivery and the exhaust pressures, respectively, and $x_s(t)$ the position of the spool servo-valve.

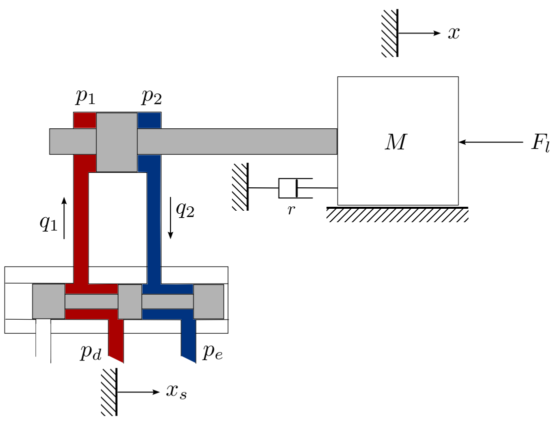


Figure 16.2: Mechanical load driven by an electro-hydraulic actuator.

Consider the following characteristic for the servo-valve

$$q(t) = k\frac{x_s(t)}{\sqrt{\rho}}\sqrt{\Delta p(t)}$$

and specifically

$$q_1(t) = k\frac{x_s(t)}{\sqrt{\rho}}\begin{cases}\sqrt{p_d - p_1(t)} & x_s(t) > 0\\ \sqrt{p_e - p_1(t)} & x_s(t) < 0\end{cases}$$

and

$$q_2(t) = k\frac{x_s(t)}{\sqrt{\rho}}\begin{cases}\sqrt{p_2(t) - p_e} & x_s(t) > 0\\ \sqrt{p_d - p_2(t)} & x_s(t) < 0\end{cases}$$

where ρ is the nominal density of the hydraulic medium, k is a sizing parameter. Then, from the continuity equations of pressure, for the two chambers we have

$$\begin{aligned}\frac{d}{dt}p_1(t) &= \frac{\beta}{V_0 + Sx(t)}\left[q_1(t) - S\frac{d}{dt}x(t) - C_L(p_1(t) - p_2(t))\right]\\ \frac{d}{dt}p_2(t) &= \frac{\beta}{V_0 - Sx(t)}\left[S\frac{d}{dt}x(t) - q_2(t) + C_L(p_1(t) - p_2(t))\right]\end{aligned} \tag{16.1}$$

where V_0 is the volume of each chamber for $x = 0$, β is the isothermal bulk modulus of the hydraulic fluid, S is the effective bore of the cylinder, and finally C_L is the leakage coefficient between the two chambers of the cylinder.

Define $p_m(t) = p_1(t) - p_2(t)$, the pressure difference between the two chambers, and assuming nearly constant and equivalent flow rates $q_1(t) = q_2(t) = q_m(t)$. Assume $p_e = 0$, then for small variations around the equilibrium point characterised by $p_m(t) = 0, \forall t$ and $x(t) \approx 0, \forall t$, the corresponding behaviour can be characterised by the following equation

$$\frac{d}{dt}p_m(t) = -\frac{2\beta C_L}{V_0}p_m(t) - \frac{2S\beta}{V_0}\frac{d}{dt}x(t) + \frac{2\beta}{V_0}q_m(t)$$

which represents, in terms of a LTI system, the dynamics of the hydraulic cylinder, where

$$q_m(t) = k\frac{x_s(t)}{\sqrt{\rho}}\sqrt{p_d}$$

Overall, the dynamics of the hydraulic cylinder can be written in terms of the following equation

$$\frac{d}{dt}p_m(t) = -\frac{2\beta C_L}{V_0}p_m(t) - \frac{2S\beta}{V_0}\frac{d}{dt}x(t) + \frac{2\beta k\sqrt{p_d}}{V_0\sqrt{\rho}}x_s(t) \tag{16.2}$$

16.1.2 Modelling the mechanical dynamics

The modelling of the mechanical load is actually straightforward if we consider the mass M of Fig. 16.2 representative of all moving masses and results

$$\begin{aligned}\frac{d}{dt}x(t) &= v(t)\\ \frac{d}{dt}v(t) &= -\frac{r}{M}v(t) + \frac{1}{M}(F(t) - F_l(t))\end{aligned} \tag{16.3}$$

where $F(t)$ represents the force produced by the cylinder, i.e.

$$F(t) = S(p_1(t) - p_2(t)) = Sp_m(t)$$

and $F_l(t)$ a generic external force (see again Fig. 16.2).

16.1.3 Control strategy

The two sets of equations (16.2) and (16.3) constitute the overall model to be used as a reference for the tuning of the control law. The reported equations correspond to the block diagram reported in Fig. 16.3.

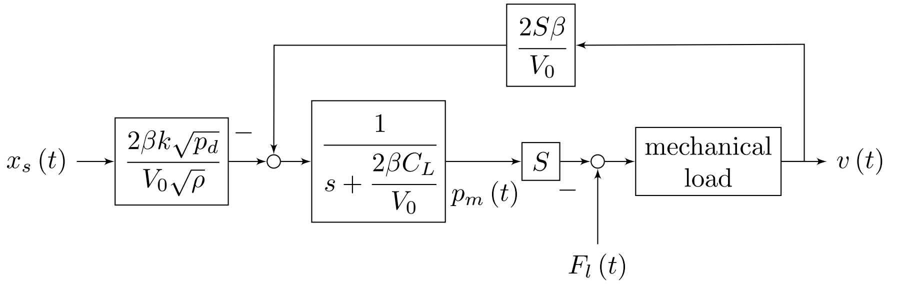


Figure 16.3: Block diagram of the overall system including both the electro-hydraulic and the mechanical dynamics.

From a control perspective, assuming the valve to be servo-actuated, one can assume the position of its spool very closed to its reference, i.e. $x_s(t) = x_s^0(t), \forall t$. By combining the two sets of equations (16.2) and (16.3), one obtains the following third-order LTI system

$$\begin{aligned}
\frac{d}{dt}p_m(t) &= -\frac{2\beta C_L}{V_0}p_m(t) - \frac{2S\beta}{V_0}v(t) + \frac{2\beta k\sqrt{p_d}}{V_0\sqrt{\rho}}x_s^0(t) \\
\frac{d}{dt}x(t) &= v(t) \\
\frac{d}{dt}v(t) &= -\frac{r}{M}v(t) + \frac{S}{M}p_m(t) - \frac{1}{M}F_l(t)
\end{aligned} \tag{16.4}$$

where $x_s^0(t)$, i.e. the position of the spool, can be regarded as an input, together with the disturbance force $F_l(t)$ applied to the load, while the state is

$$\boldsymbol{x}(t) = \begin{bmatrix} p_m(t) & x(t) & v(t) \end{bmatrix}^T$$

Further inspecting the equations, one can notice that the system corresponds to the following characteristic polynomial

$$\det(s\boldsymbol{I} - \boldsymbol{A}) = \frac{s\left(2\beta\left(S^2 + rC_L\right) + \left(rV_0 + 2\beta C_L M\right)s + MV_0 s^2\right)}{MV_0}$$

Assuming a negligible friction coefficient, i.e. $r \approx 0$, the characteristic polynomial further simplifies to

$$\det(s\boldsymbol{I} - \boldsymbol{A}) \approx s\left(\frac{2\beta S^2}{MV_0} + \frac{2\beta C_L}{V_0}s + s^2\right)$$

which clearly indicates the presence of an eigenvalue in the origin and two complex and conjugate ones with the following natural frequency and damping

$$\omega_n = S\sqrt{\frac{2\beta}{MV_0}}, \xi = \frac{\sqrt{2}}{2}\frac{C_L}{S}\sqrt{\frac{\beta M}{V_0}}$$

respectively. The transfer function from the reference position of the spool of the valve $x_s^0(t)$ to the position $x(t)$ of the load can be then written as

$$G(s) = \frac{\mu/s}{\frac{s^2}{\omega_n^2} + 2\frac{\xi}{\omega_n}s + 1}$$

where

$$\mu \approx \frac{k\sqrt{p_d}}{\sqrt{\rho}S}$$

16.2 Synthesis of the position controller

The selection of the controller naturally falls within the family of PID controllers. A purely proportional controller constitutes the simplest choice, though the absence of the integral action would make the steady-state compensation of constant disturbances (as e.g. the force $F_l(t)$) not possible. As a consequence, the precision in the positioning of the rod of the cylinder will be demanded a high value for the generalised gain of the loop transfer function $K_p\mu$.

Example 80. *[Tuning of the position controller] Consider a generic control surface of an aircraft as the one reported in Fig. 16.4. For the system at hands, we assume the following*

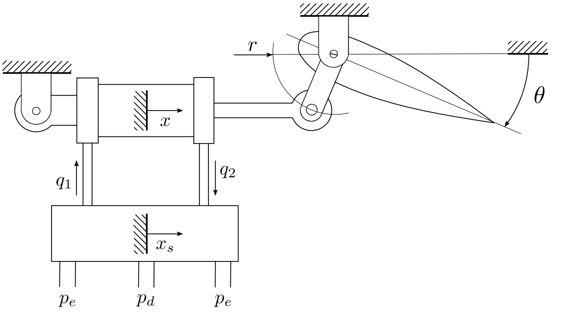


Figure 16.4: Control surface with an electro-hydraulic actuator.

values for the relevant parameters. For the hydraulic fluid $\beta = 1.5\ GPa$, $\rho = 900\ kg/m^3$, for the cylinder $V_0 = 6.2832 \cdot 10^{-5}\ m^3$, $S = 3.1416 \cdot 10^{-4}\ m^2$, $C_L = 1.4 \cdot 10^{-11}\ m^4 s/kg$, for the valve $k = 2.7 \cdot 10^{-4}$ and $p_d = 21\ MPa$. Finally for the mechanical load, we assume $m = 0.5$ kg the mass of the piston, $J = 0.16\ kg \cdot m^2$ the moment of inertia of the control surface, and

$r = 10$ cm the radius of the last arm of the kinematic chain. For small displacements of the rod of the cylinder, we can approximate

$$\theta(t) \approx -\frac{x(t)}{r}$$

hence, the overall reflected mass can be regarded as

$$M = m + \frac{J}{r^2} = 16.5 \ kg$$

Consistently, the transfer function from the position of the spool to the displacement of the rod of the cylinder results as follows

$$G(s) = \frac{3.749 \cdot 10^7}{s\left(s^2 + 668.5s + 2.856 \cdot 10^5\right)}$$

corresponding to $\omega_n \approx 534 \ rad/s$, $\xi \approx 0.625$, and $\mu \approx 131$.

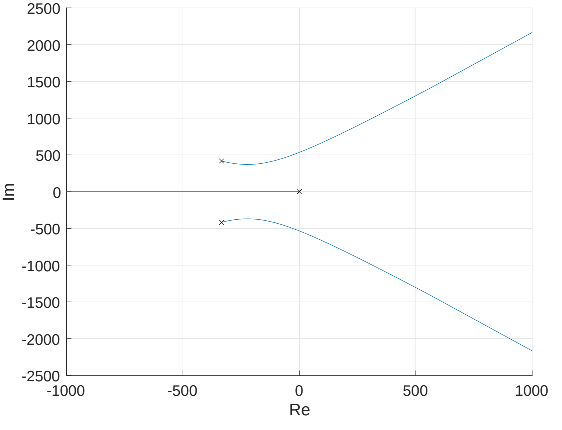


Figure 16.5: Root locus for varying proportional gains K_p.

The root locus of the corresponding loop transfer function $L(s) = K_p G(s)$ is reported in Fig. 16.5. As one can see, in order to guarantee asymptotic stability, there is an upper limit for the proportional gain. This limit, say $K_p^{\max}$ can be easily computed by applying the conservation of the sum, which applies since the relative degree ν of the loop transfer function is $\nu = 3 \geq 2$. Specifically, when the two closed-loop roots are on the imaginary axis, the other one corresponds to the sum of all open-loop poles, and specifically

$$s = 0 - 2\xi\omega_n = -668.4508$$

The corresponding maximum value for the proportional gain is therefore

$$K_p^{\max} = \frac{1}{|G(-668)|} = 5.0918$$

In order to stay sufficiently far from the stability limit of the system happening for $K_p = K_p^{\max}$, we can set $K_p = 0.1 K_p^{\max} = 0.5092$. Correspondingly, the crossover frequency results as follows

$$\omega_c = K_p \mu = 0.1 K_p^{\max} \mu \approx 67 \ rad/s$$

which is sufficiently far from the resonant behaviour of the system at $\omega_n = 534 \ rad/s$. The response of the angular displacement of the control surface $\theta(t)$ for a step change in its reference position $\theta^0(t) = 10$ deg, $t \geq 0$ is reported in Fig. 16.6.

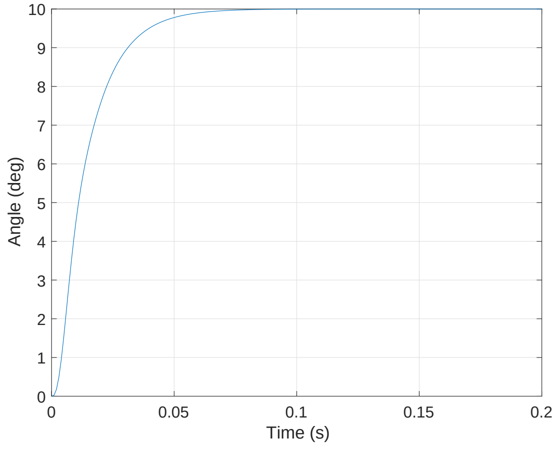


Figure 16.6: Step response of the angular displacement of the control surface $\theta(t)$ for a proportional controller tuned with gain $K_p = 0.1K_p^{\max}$.

Chapter 17

Motion control with electro-mechanical actuators

The problem of controlling the motion of a mechanical system actuated by an electrical motor also entails a series of methodological and technological challenges. This Chapter is then intended to address all these aspects for this specific case of adoption of feedback control.

17.1 Components for motion control

In the following we will make reference to a very simple mechanism composed by an electrical motor, together with its driver, a gear-box, and a generic load. The position of the rotating axis will be measured by a proper (typically optical) sensor.

17.1.1 Modelling the electrical dynamics

Without any lack of generality, we will make reference to a permanent magnet DC (PMDC) motor. First, consider the working principle of a DC motor: the voltage applied to rotor windings generates a current, the windings which are current-currying conductors, are immersed in the stator magnetic field. Lorenz forces are then exerted on the windings and produce motion of the rotor. As the rotor rotates, the variable flux through the windings produces Lenz voltage (called **back electromotive force, back-EMF**) which opposes to the supplied voltage.

From a purely electrical perspective, a PMDC motor corresponds to the electrical circuit in Fig. 17.1 where $v(t)$ is the supplied voltage, $e(t)$ is the back-EMF, while R and L are the resistance and the inductance of the rotor windings, respectively.

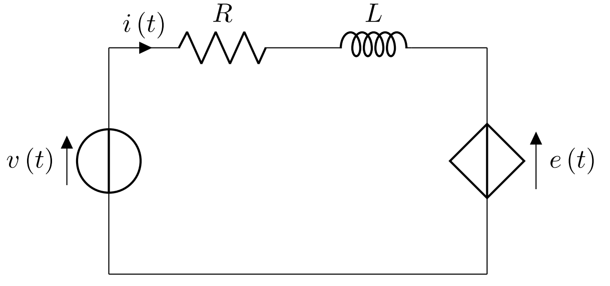


Figure 17.1: Electrical model of a PMDC motor.

Letting $i(t)$ the current flowing in the circuit, the equations governing the electrical dynamics of the motor are given in the following

$$\begin{aligned} \frac{d}{dt} i(t) &= -\frac{R}{L} i(t) + \frac{1}{L}(v(t) - e(t)) \\ e(t) &= k_e \omega_m(t) \end{aligned} \tag{17.1}$$

where the back-EMF $e(t)$ has been assumed, as it typically happens, proportional to the spinning velocity of the rotor $\omega_m(t)$, through coefficient k_e. It is easy to recognise a first-order LTI system with input the net voltage $v(t) - e(t)$ and output the current $i(t)$. Hence, from the Laplace perspective, the model of the PMDC motor can be written in the following terms

$$I(s) = \frac{1}{Ls + R}(V(s) - E(s))$$

17.1.2 Modelling the mechanical dynamics

The simplest mathematical model of the overall mechanical load can be regarded as the two-inertia system reported in Fig. 17.2, where J_m and J_l represent the moment of inertia of the rotor and of the load, respectively, $\eta \geq 1$ is the gear-box reduction ratio, $\tau_m(t)$ is the driving torque, $\omega_m(t)$ and $\omega_l(t)$ are the rotational velocities of the rotor and of the load, respectively, and finally $\tau_l(t)$ represents an additional load torque.

The equations governing the behaviour of the mechanical quantities are

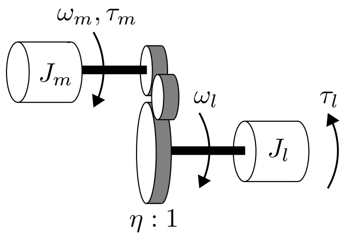


Figure 17.2: Mechanical model.

given in the following

$$
\begin{aligned}
\frac{d}{dt}\omega_m(t) &= \frac{1}{J_m}(\tau_m(t) - \tau_{lm}(t)) \\
\frac{d}{dt}\omega_l(t) &= \frac{1}{J_l}(\eta\tau_{lm}(t) - \tau_l(t)) \\
\tau_m(t) &= k_t i(t)
\end{aligned}
\tag{17.2}
$$

where the driving torque $\tau_m(t)$ has been assumed proportional to the current $i(t)$ flowing in the rotor windings through constant k_t, and $\eta\tau_{lm}$ represents the driving torque to the load.

A gear-box is typically placed between the motor and the mechanical load to operate. In mechatronic system governed by electrical motors, fixed-ratio gear-boxes are usually adopted, hence the reduction ratio $n \geq 1$ will be considered constant. As one can see from (17.2), the equations governing the mechanical quantities of the system is incomplete since a relationship between rotor- and load-side quantities is missing. The two typical possibilities for the modelling of the transmission stage will be introduced in the following.

17.1.3 Control strategy

Equations (17.1) and (17.2) together constitute the overall electro-mechanical model to be used as a reference for the tuning of the control law. The two last algebraic equations, i.e. $e(t) = k_e\omega_m(t)$ and $\tau_m(t) = k_t i(t)$, represents the mutual coupling between mechanical and electrical quantities.
Regardless how the remaining modelling issue is solved, the block diagram of the overall system is reported in Fig. 17.3.

From a control perspective, the voltage $v(t)$ applied to rotor windings represents the control variable, while the current $i(t)$, the rotor velocity $\omega_m(s)$, and its integral, i.e. the rotor position $q_m(t)$, are variables to be controlled or monitored. From now on, we will consider the most complete scenario in which both the velocity $\omega_m(t)$ and the position $q_m(t)$ of the rotor shaft have to be

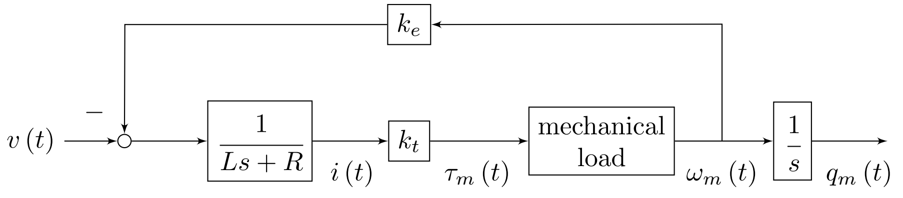


Figure 17.3: Block diagram of the overall system including both the electrical and the mechanical dynamics.

controlled. Motion control of lathes, for which only the spinning velocity has to be controlled, will not fall outside the forthcoming discussions, as we will see.

Due to the availability of three measurements, $i(t)$, $\omega_m(s)$, and $q_m(t)$, and their relative difference in terms of frequency components, we can easily arrange a three-level cascaded control scheme with an inner control loop dedicated to the current with controller to be designed $R_i(s)$, an intermediate one for the velocity with controller $R_v(s)$, and the outermost for the position controlled by $R_p(s)$. The corresponding block diagram is reported in Fig. 17.4.

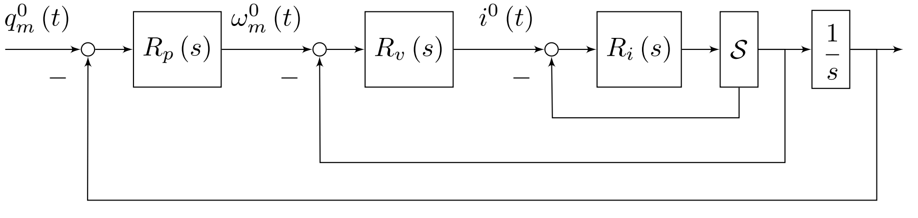


Figure 17.4: Typical control architecture for motion control problems with three nested (cascaded) control loops.

17.2 Synthesis of the current controller

In this Section, design guidelines for the tuning of the current controller $R_i(s)$ of Fig. 17.4 will be derived. Looking at the schematic representation of the electro-mechanical dynamics in Fig. 17.3, the model of the system to be controlled has the input voltage $v(t)$ as input and the current $i(t)$ flowing in the rotor windings as corresponding output, and therefore

$$I(s) = \frac{\frac{1}{Ls+R}}{1+\frac{1}{Ls+R}k_e k_t G_v(s)}V(s) = \left(\frac{Ls+R}{Ls+R+k_e k_t G_v(s)}\right)\frac{1}{Ls+R}V(s)$$

where $G_v(s)$ represents the transfer function from the torque $\tau_m(t)$ to the velocity of the rotor $\omega_m(t)$. Notice that the term within parentheses is a term characterised by extremely low-frequency components with respect to those of the electrical dynamics and will be regarded as a low-frequency uncertainty for our synthesis purposes. Moreover, since reasonably $G_v(s) = 0$ for $s \to +\infty$, we have

$$\lim_{s\to+\infty} \frac{Ls+R}{Ls+R+k_e k_t G_v(s)} = \lim_{s\to+\infty} \frac{Ls}{Ls+k_e k_t G_v(s)} = 1$$

As a consequence of the discussion on the sensitivity transfer function reported in Chapter 7, one can focus on the following relationship

$$I(s) \approx \frac{1}{Ls+R} V(s)$$

for the design of the current controller $R_i(s)$. Being the system to be controlled characterised by a single left half-plane pole and no zeros, a proportional-integral controller could be regarded as the best choice, hence

$$R_i(s) = K_{pi} \frac{1+sT_{ii}}{sT_{ii}}$$

In order to achieve a crossover frequency of ω_{ci}, the zero of the controller can be placed approximately one decade before ω_{ci}, i.e. $T_{ii} = 10/\omega_{ci}$. This way, in the neighbourhood of the crossover frequency ω_{ci}, the corresponding loop transfer function would result as follows

$$L_i(s) \approx K_{pi} \frac{1}{Ls+R} \frac{1+sT_{ii}}{sT_{ii}} \approx \frac{K_{pi}}{Ls}$$

from which the following tuning for the proportional gain K_{pi} can be obtained

$$K_{pi} = L\omega_{ci}$$

Example 81. *[Tuning of the current controller] Consider the data-sheet of a commercial PMDC motor in Fig. 17.5.*

Assuming one wants to tune the current controller for the specific model C23-L33-10, the parameters of interest are the rotor (named terminal in the data-sheet) resistance $R = 0.6$ Ω *and inductance* $L = 0.35\ mH$*. Notice that the data-sheet also reports the electrical time constant which can be defined as* $\tau_e = L/R = 0.5833\ ms$*. The open-loop system has its pole at frequency* $\omega_e = 1/\tau_e = R/L \approx 1700\ rad/s$*, we than assume* $\omega_{ci} = 8000\ rad/s$ *to be a reasonable requirement for the closed-loop system. Therefore, the proportional-integral current controller* $R_i(s)$ *can be tuned using the following parameters*

$$T_{ii} = 10/\omega_{ci} = 10/8000 = 13\ ms, K_{pi} = L\omega_{ci} = 2.8\ V/A$$

The corresponding phase margin results of $\psi_m = 96.5$ deg*. The step responses for a reference current of* $i^0(t) = 0.1\ A$ *is reported in Fig. 17.6.*

C23 Series Specifications

Part Number*		C23-L33					C23-L40				
Winding Code**		10	20	30	40	50	10	20	30	40	50
L = Length	inches	3.33					4				
	millimeters	84.6					101.6				
Terminal Voltage	volts DC	12	12	24	36	48	12	24	36	48	60
Rated Speed	RPM	4700	2150	4200	3750	3000	2300	3600	3500	2850	2250
	rad/sec	492	225	440	393	314	241	377	367	298	236
Rated Torque	oz-in	7.5	12.6	12.7	14.4	15.8	17.3	25.5	25.3	25.6	24.2
	Nm	0.05	0.09	0.09	0.10	0.11	0.12	0.18	0.18	0.18	0.17
Rated Current	Amps	4.75	4.3	3	2	1.4	4.9	4.3	2.75	1.8	1.1
Torque Senstivity	oz-in/amp	2.65	4.25	6.2	10.25	15.75	4.84	7.74	12	18.5	28.75
	Nm/amp	0.0187	0.0300	0.0438	0.0724	0.1112	0.0342	0.0547	0.0847	0.1306	0.2030
Back EMF	volts/KRPM	2	3.15	4.6	7.6	11.5	3.58	5.72	8.82	13.82	21.22
	volts/rad/sec	0.0191	0.0301	0.0439	0.0726	0.1098	0.0342	0.0546	0.0842	0.1320	0.2026
Terminal Resistance	ohms	0.60	1.00	1.70	4.00	9.00	0.70	0.96	2.30	5.50	12.00
Terminal Inductance	mH	0.35	0.94	2.00	5.50	13.00	0.50	1.30	3.10	7.36	18.00
Rotor Inertia	oz-in-sec^2	0.0022	0.0022	0.0022	0.0022	0.0022	0.004	0.004	0.004	0.004	0.004
	g-cm^2	155.4	155.4	155.4	155.4	155.4	282.5	282.5	282.5	282.5	282.5
Electrical Time Constant	millisecond	0.5833	0.9400	1.1765	1.3750	1.4444	0.7143	1.3584	1.3478	1.3382	1.5000

Figure 17.5: Excerpt of the data-sheet of a commercial PMDC electrical motor (Moog).

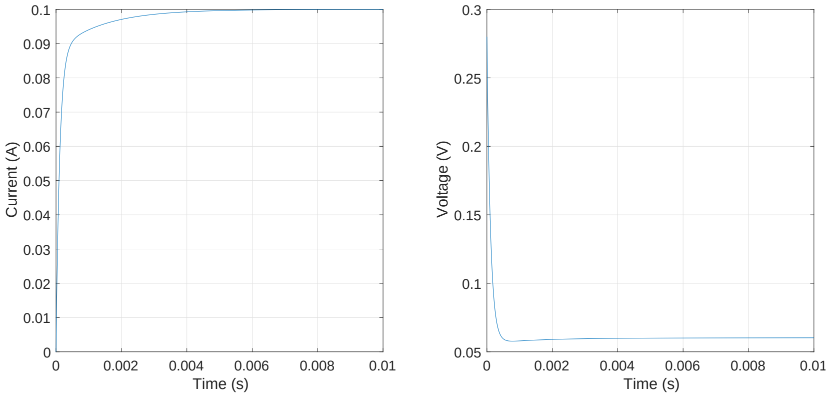


Figure 17.6: Step responses of the current (left) and of the supplied voltage (right) for a reference current of 0.1 A.

17.3 Synthesis of the velocity controller

Consider again the mechanical system in Fig. 17.2. As already discussed, its mathematical model in (17.2) lacks of modelling details regarding the relationship between motor- and load-side quantities. In the following, we will address two main modelling assumptions, that is

1. the transmission stage can be considered as infinitely rigid;
2. the transmission stage involve some non-negligible flexibilities that have to be taken into account.

17.3.1 Rigid case

The case of a completely rigid transmission stage actually constitutes a very strong assumption. In this setting, motor- and load-side quantities can be regarded as kinematically related by the following algebraic relationships:

$$q_l(t) = \frac{q_m(t)}{\eta}, \omega_l(t) = \frac{\omega_m(t)}{\eta}$$

From these relationships, once substituted in (17.2), one can easily derive a lump model which simultaneously describes both the motor- and the load-side dynamical behaviours, i.e.

$$\begin{aligned} \frac{d}{dt}\omega_m(t) &= \frac{1}{J_m + \dfrac{J_l}{\eta^2}}\left(\tau_m(t) - \frac{\tau_l(t)}{\eta}\right) \\ \tau_m(t) &= k_t i(t) \\ \omega_l(t) &= \frac{\omega_m(t)}{\eta} \end{aligned} \tag{17.3}$$

The corresponding relationship between the current $i(t)$ and the rotor velocity $\omega_m(t)$, in the Laplace domain, can be written as follows

$$\Omega_m(s) = \frac{1}{\left(J_m + \dfrac{J_l}{\eta^2}\right)s}\left(k_t I(s) - \frac{T_l(s)}{\eta}\right)$$

which clearly highlights an integral dynamics between the net torque $k_t i(t) - \tau_l(t)/\eta$ and the motor velocity $\omega_m(t)$. For this reason, a proportional-integral controller still constitutes a reasonable choice for the velocity controller $R_v(s)$. Notice that, in order to nullify the steady-state error in case of a constant load disturbance torque $\tau_l(t) = \tau_l, \forall t \geq 0$, the integral action within the controller $R_v(s)$ is needed. In fact

$$e_\infty = \lim_{s\to 0} sE(s) = -\lim_{s\to 0} \frac{\tau_l/\eta}{\left(J_m + \dfrac{J_l}{\eta^2}\right)s + R_v(s)}$$

This said, assuming the motor current to be ideally regulated by the inner current control loop, i.e. $i(t) = i^0(t)$, the loop transfer function to be considered for the tuning of the velocity controller $R_v(s)$ results as follows

$$L_v(s) = K_{pv}\frac{1 + sT_{iv}}{sT_{iv}}\frac{k_t}{\left(J_m + \dfrac{J_l}{\eta^2}\right)s}$$

Hence, given the design specification in terms of a desired crossover frequency $\omega_{cv} \ll \omega_{ci}$, one can place the zero of the controller one decade before, i.e. $T_{iv} = 10/\omega_{cv}$, and set the proportional gain K_{pv} as

$$K_{pv} = \left(J_m + \frac{J_l}{\eta^2}\right)\frac{\omega_{cv}}{k_t}$$

so to obtain the desired crossover frequency.

Example 82. *[Tuning of the velocity controller] Consider again the C23-L33-10 PMDC motor of Fig. 17.5 with a load $J_l = 15.54\ kg \cdot m^2$ and a transmission ratio of $\eta = 100$. The parameters of interest are the rotor inertia $J_m = 155.4\ g \cdot cm^2$, and the torque constant (torque sensitivity) $k_t = 0.0187\ Nm/A$. Assuming $\omega_{cv} = 500\ rad/s$ to be a good separation between the inner current loop and the velocity control loop to be tuned, a proportional-integral current controller $R_v(s)$ can be tuned using the following parameters*

$$T_{iv} = 10/500 = 12.5\ ms, K_{pv} = \left(J_m + \frac{J_l}{\eta^2}\right)\frac{\omega_{cv}}{k_t} = 133\ As/rad$$

The corresponding phase margin results of $\psi_m = 84.3$ deg. Bode's diagrams of the velocity loop transfer function, obtained by neglecting the presence of the inner current control loop, are given in Fig. 17.7.

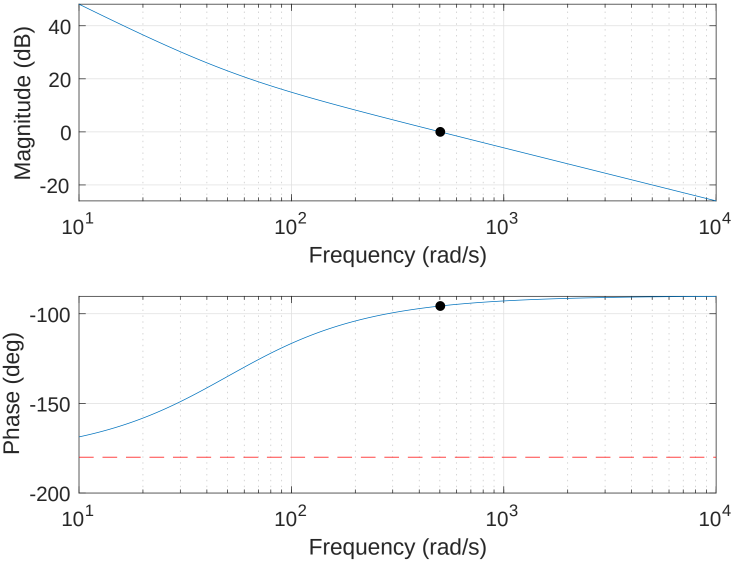
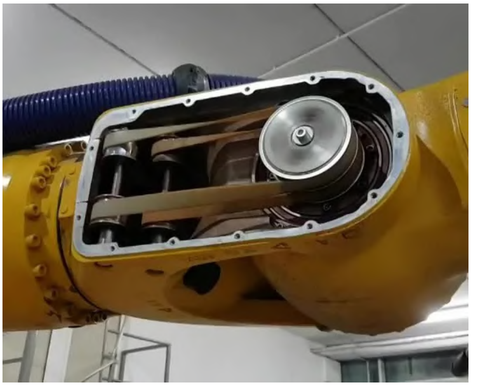


Figure 17.7: Bode's diagrams of the velocity loop.

17.3.2 Elastic case

The assumption of a completely rigid transmission stage is clearly too optimistic. As seen for the tuning of the velocity controller $R_v(s)$, there is apparently no practical limitation, apart from the presence of high-frequency measurement noise and the presence of the inner current loop, in the selection of the crossover frequency ω_{cv}. As a matter of fact in many commercial machines, the transmission stage cannot be considered as completely rigid (see Fig. 17.8 for a few examples) being manufactured of deformable materials.

(a) Transmission between the motor and the drum of a washing machine (Creative Commons)



(b) Transmission within the wrist of an industrial robot

Figure 17.8: Example of transmission stages.

It follows that the modelling assumptions adopted so far should be retracted. A more realistic model of the transmission stage can be derived assume the transmission as a spring-damper system between the motor- and the load-sides. Correspondingly, the driving torque τ_{lm} in (17.2) can be written as follows:

$$\tau_{lm}(t) = k_{el}(q_m(t) - \eta q_l(t)) + d_{el}(\omega_m(t) - \eta\omega_l(t))$$

where parameters k_{el} and d_{el} represent the stiffness and the damping of the transmission elements, respectively. Notice that, in case of belt transmissions, while the stiffness coefficient can be somehow estimated from the Young's modulus E of the material, the area A, and the nominal length L_0, i.e.

$$k_{el} = \frac{EA}{L_0}$$

an estimate for the damping coefficient d_{el} is definitely more difficult to be obtained without experimental data.

Using the elastic model of the transmission just introduced, starting from (17.2), the equations governing the behaviour of the mechanical system are given in the following:

$$\begin{aligned} \frac{d}{dt} q_m(t) &= \omega_m(t) \\ \frac{d}{dt} \omega_m(t) &= \frac{1}{J_m}\left(k_t i(t) - \tau_{lm}(t)\right) \\ \frac{d}{dt} \eta q_l(t) = \eta \frac{d}{dt} q_l(t) &= \eta \omega_l(t) \\ \frac{d}{dt} \eta \omega_l(t) = \eta \frac{d}{dt} \omega_l(t) &= \frac{\eta}{J_l}\left(\eta \tau_{lm}(t) - \tau_l(t)\right) \\ \tau_{lm}(t) &= k_{el}\left(q_m(t) - \eta q_l(t)\right) + d_{el}\left(\omega_m(t) - \eta \omega_l(t)\right) \end{aligned} \tag{17.4}$$

which is a LTI dynamical system of order $n = 4$ and state vector

$$\boldsymbol{x}(t) = \begin{bmatrix} q_m(t) & \omega_m(t) & \eta q_l(t) & \eta \omega_l(t) \end{bmatrix}^T$$

The equations in (17.4) correspond to the following relationship between the input current $i(t)$ and the rotor velocity $\omega_m(t)$:

$$\Omega_m(s) = \frac{k_t}{s\left(J_m + \dfrac{J_l}{\eta^2}\right)} \frac{1 + 2\dfrac{\xi_z}{\omega_z}s + \dfrac{s^2}{\omega_z^2}}{1 + 2\dfrac{\xi_p}{\omega_p}s + \dfrac{s^2}{\omega_p^2}} I(s) \tag{17.5}$$

where

$$\omega_z = \sqrt{\frac{\eta^2 k_{el}}{J_l}}, \omega_p = \omega_z\sqrt{1+\rho}, \xi_z = \frac{d_{el}}{2}\frac{\eta}{\sqrt{J_l k_{el}}}, \xi_p = \xi_z\sqrt{1+\rho}, \rho = \frac{J_l}{\eta^2 J_m}$$

The transfer function in (17.5) is characterised by a pole in the origin, and two pairs of typically poorly damped complex and conjugate poles and zeros. Moreover, since $\rho > 0$ and therefore $\sqrt{1+\rho} > 1$, the frequency of the complex and conjugate zeros ω_z anticipates the one of the complex and conjugate poles ω_p, i.e. $\omega_z < \omega_p$. Finally notice that the low-frequency behaviour of the transfer function, i.e.

$$\frac{k_t}{s\left(J_m + \dfrac{J_l}{\eta^2}\right)}$$

is identical to the one derived under the assumption of an infinitely rigid transmission stage.

Example 83. *Considering the data in the last example with* $\omega_z = 500\ rad/s$ *and* $\xi_z = 0.1$, *the Bode's diagrams of the frequency response of the transfer function in* (17.5) *as compared to its low-frequency (rigid) approximation are given in Fig. 17.9.*

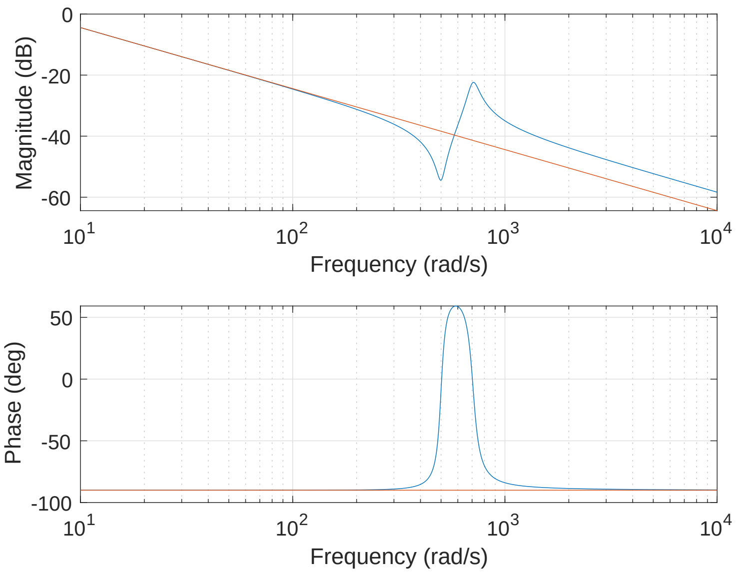


Figure 17.9: Bode's diagrams of the frequency response of the transfer function relating the input current $i(t)$ and the motor velocity $\omega_m(t)$ (blue) and its low-frequency approximation (red).

Adopting the same tuning for the proportional-integral controller derived for the rigid case, one obtains the step response of the motor velocity $\omega_m(t)$ reported in Fig. 17.10 together with the corresponding scaled load velocity $\eta\omega_l(t)$.

Though the closed-loop remains asymptotically stable, evident oscillations can be appreciated both the rotor velocity $\omega_m(t)$ and in the scaled load velocity $\eta\omega_l(t)$.

In the last example, we have seen that, though stability is not compromised, in case of a non infinitely rigid transmission stage, significant oscillations might arise if the proportional-integral velocity controller $R_v(s)$ is not properly tuned. Consider then the loop transfer function for the tuning of the velocity control law, i.e.

$$L_v(s) = K_{pv}\frac{1+sT_{iv}}{sT_{iv}}\frac{k_t}{s\left(J_m+\dfrac{J_l}{\eta^2}\right)}\frac{1+2\dfrac{\xi_z}{\omega_z}s+\dfrac{s^2}{\omega_z^2}}{1+2\dfrac{\xi_p}{\omega_p}s+\dfrac{s^2}{\omega_p^2}}$$

If one assume the zero of the controller to be placed approximately one decade before the frequency of the complex and conjugate zero, i.e. $T_{iv} = 10/\omega_z$, the

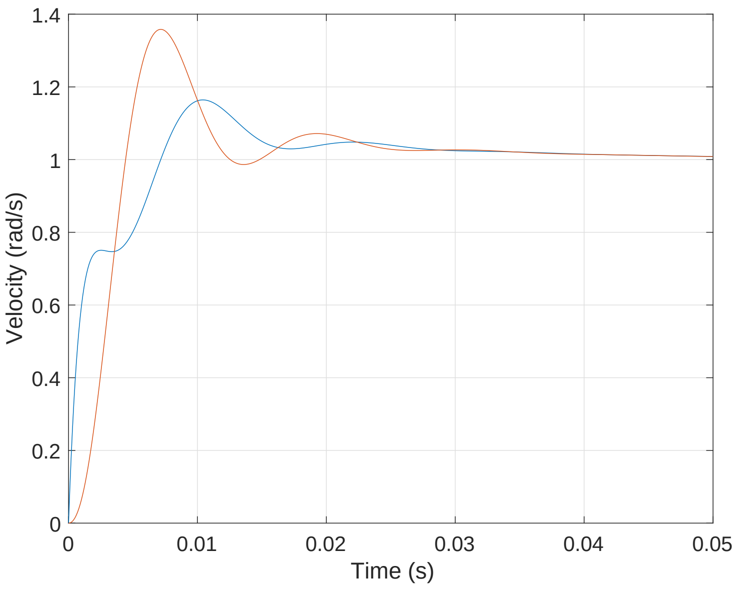


Figure 17.10: Closed-loop unit step response of the motor velocity $\omega_m(t)$ (blue) and corresponding scaled load velocity $\eta\omega_l(t)$ (red).

loop transfer function depends uniquely on the tunable parameter K_{pv}, in fact

$$L_v(s) = K_{pv}\frac{1+10s\omega_z}{s10/\omega_z}\frac{k_t}{s\left(J_m+\dfrac{J_l}{\eta^2}\right)}\frac{1+2\dfrac{\xi_z}{\omega_z}s+\dfrac{s^2}{\omega_z^2}}{1+2\dfrac{\xi_p}{\omega_p}s+\dfrac{s^2}{\omega_p^2}}$$

The root locus applied to the last transfer function by varying K_{pv} might certainly help in understanding whether a trade-off between promptness of the response and absence of oscillations can be found. For the parameters adopted in the last example, the root locus applied to $L_v(s)$ and evaluated for varying K_{pv} is reported in Fig. 17.11. The root locus of Fig. 17.11 reveals many (rather obvious) aspects:

- for **low values of** K_{pv}, the closed-loop complex and conjugate poles are poorly damped, as the **open-loop complex and conjugate poles** are also purely damped;
- for **high values of** K_{pv}, the closed-loop complex and conjugate poles are also expected to be poorly damped, as **open-loop complex and conjugate zeros** are poorly damped.

The reason for the first aspect is that for small values of K_{pv}, closed-loop roots are still close to the corresponding open-loop positions. The reason for the

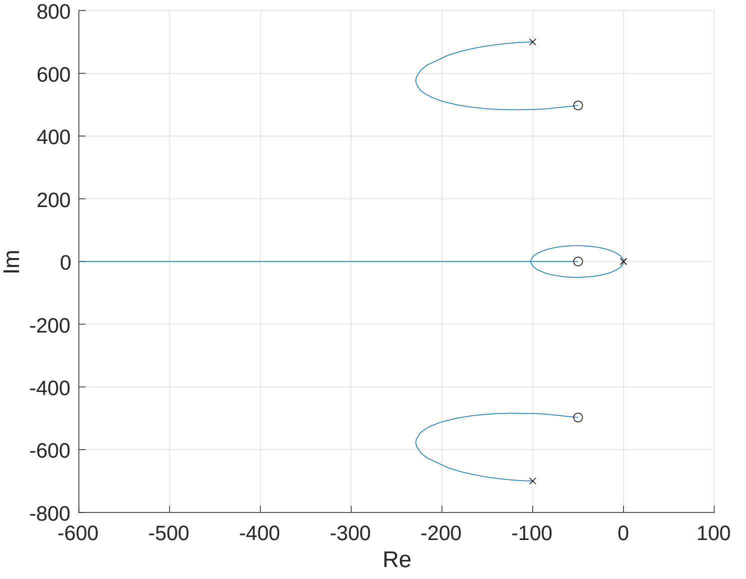


Figure 17.11: Root locus for the loop transfer function $L_v(s)$ for the case $T_{iv} = 10/\omega_z$.

latter observation is than for high gains, closed-loop poles approach open-loop zeros, and the transfer function in (17.5) has indeed poorly damped zeros. Therefore, the best trade-off between promptness of the response and possible absence of closed-loop oscillations both at the motor- and the load-side is to select the value for the proportional gain K_{pv} that **maximises the damping ξ_v of closed-loop complex and conjugate poles**.

In order to understand how to select the optimal proportional gain K_{pv} that maximises the closed-loop damping, we need to address the behaviour of ξ_v as a function of the proportional gain K_{pv}. For generality, this behaviour is represented in Fig. 17.12 in terms of the corresponding crossover frequency $\tilde{\omega}_{cv}$ as evaluated on the rigid model and normalised by ω_z, i.e.

$$\tilde{\omega}_{cv} = \frac{1}{\omega_z}\frac{k_t K_{pv}}{J_m + \dfrac{J_l}{\eta^2}}$$

As expected from the corresponding root locus, the damping ξ_v first increases with $\tilde{\omega}_{cv}$ and then decreases. The maximum value of the closed-loop damping $\xi_v^{\max}$ is obtained for $\tilde{\omega}_{cv} \approx 0.8$, i.e. for

$$K_{pv} = \left(J_m + \frac{J_l}{\eta^2}\right)\frac{0.8\omega_z}{k_t}$$

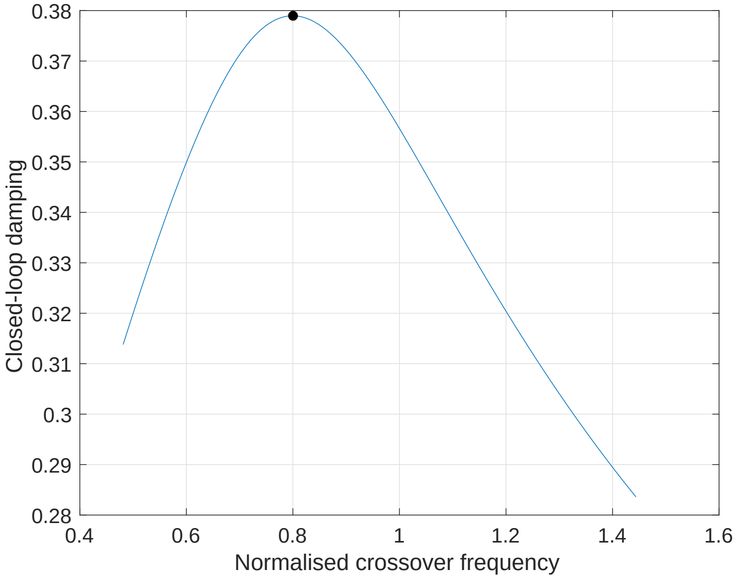


Figure 17.12: Behaviour of the closed-loop damping ξ_v as a function of $\tilde{\omega}_{cv}$ and the corresponding maximum damping $\xi_v^{\max}$.

meaning that the best tuning is to consider the rigid approximation of the model and to set $\omega_{cv} = 0.8\omega_z$.

Example 84. *Considering again the data in the last example with $\omega_z = 500$ rad/s. The tuning of the proportional-integral controller that maximises the damping ξ_v of the closed-loop poles is obtained by setting*

$$T_{iv} = 10/500 = 12.5\ ms, K_{pv} = \left(J_m + \frac{J_l}{\eta^2}\right)\frac{0.8\omega_z}{k_t} = 66.5\ As/rad$$

The corresponding closed-loop step responses for the motor velocity $\omega_m(t)$ and the scaled load velocity $\eta\omega_l(t)$ are reported in Fig. 17.13. A state-space controller can be regarded as an alternative to the proportional-integral velocity controller $R_v(s)$ thanks to its capability to arbitrarily position closed-loop roots. Assume then a zero-error controller to be designed so that the eigenvalues of the closed loop system are

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}_e - \boldsymbol{B}_e\boldsymbol{K}_e) = (\lambda - 0.8\omega_z)^4$$

and

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A} - \boldsymbol{L}\boldsymbol{C}) = (\lambda - 3000\omega_z)^3$$

The following controller would be obtained:

$$R_v(s) = 10^6\frac{-3.758s^3 + 1.111\cdot 10^3 s^2 + 2.435\cdot 10^6 s + 2.298\cdot 10^8}{s^4 + 1.04\cdot 10^4 s^3 + 8.5\cdot 10^7 s^2 + 4.805\cdot 10^{10} s}$$

The corresponding step responses for both the rotor and the load velocities are given in Fig. 17.14.

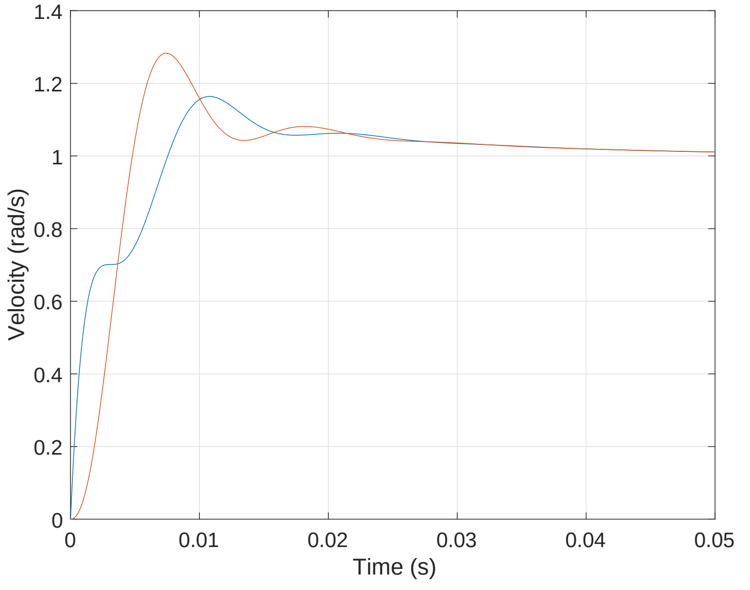


Figure 17.13: Closed-loop unit step response of the motor velocity $\omega_m(t)$ (blue) and corresponding scaled load velocity $\eta\omega_l(t)$ (red) for the optimal tuning of the velocity controller which maximises the closed-loop damping.

Despite the controller has been designed to achieve real closed-loop eigenvalues, no substantial benefits can be appreciated within the corresponding responses. Moreover, for the given specification, the transfer function of the controller $R_v(s)$ results a non-minimum phase one.

For completeness, Fig. 17.15 reports the magnitude diagram of the frequency response of the complementary sensitivity of the velocity loop $F_v(s)$ corresponding to the controller $R_v(s)$ designed with state-space methods. Thought the poles of $F_v(s)$ have been designed to be real, the magnitude diagram reports the presence of the poorly damped complex and conjugate zeros at ω_z. This should not surprise the reader, since

$$F_v(s) = \frac{L_v(s)}{1 + L_v(s)}$$

and no matter where poles are, zeros of the closed-loop system are zeros of the loop transfer function $L_v(s)$ and then those of $G_{vm}(s)$. As a matter of facts, frequency ω_z always represent a practical limitation in the bandwidth that one can achieve with a velocity controller, regardless how it has been tuned.

17.4 Synthesis of the position controller

The tuning of the velocity controller $R_v(s)$ was actually the more involved of the three. With reference to the block diagram of Fig. 17.3, the synthesis of the position controller $R_p(s)$ can be based on the transfer function from the reference motor velocity $\omega_m^0(t)$ and the position of the rotor $q_m(t)$. Disregarding

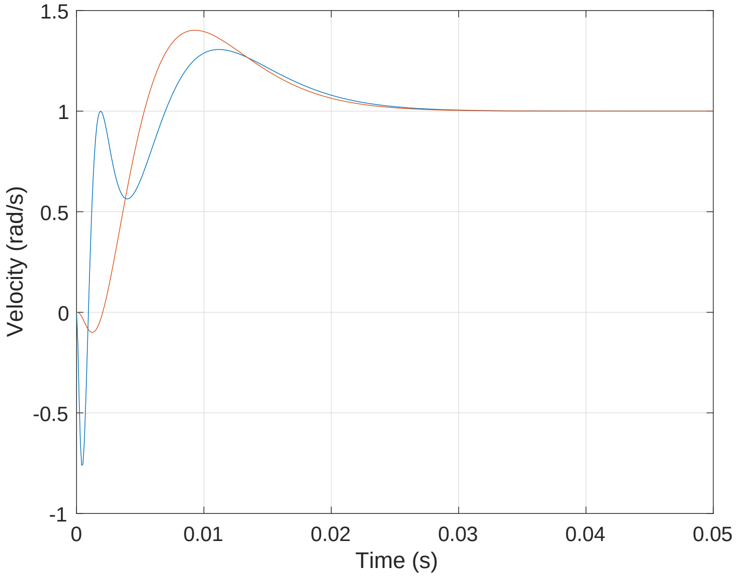


Figure 17.14: Closed-loop unit step response of the motor velocity $\omega_m(t)$ (blue) and corresponding scaled load velocity $\eta\omega_l(t)$ (red) for alternative controller designed with state-space methods.

the presence of the two inner control loops (for the current and for the velocity, respectively) which are considered as ideal, the selection of the position controller can be based on the following relationship

$$Q_m(s) = \frac{1}{s}\Omega_m(s)$$

Hence, a proportional controller with gain K_{pp} can be selected for the purpose. The corresponding loop transfer function would result as follows

$$L_p(s) = \frac{K_{pp}}{s}$$

Selecting the corresponding crossover frequency $\omega_{cp} = K_{pp}$ at least one decade before the one of the velocity controller $\omega_{cv} = 0.8\omega_z$, one obtains

$$K_{pp} = 0.08\omega_z$$

Example 85. *Still referring to the mechanical parameters already adopted in the previous examples, a possible tuning for the position controller would be the following*

$$K_{pp} = 0.08\omega_z = 40\ s^{-1}$$

The corresponding step responses of the motor and the load are reported in Fig. 17.16.

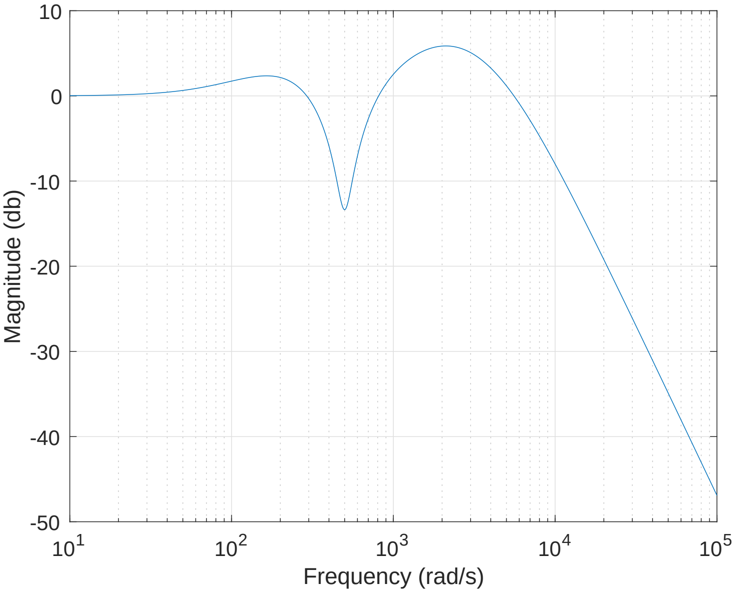


Figure 17.15: Bode's magnitude diagram of the frequency response of the complementary sensitivity for alternative controller designed with state-space methods.

Example 86. *[Control of a heavy-duty gantry crane] Consider a heavy-duty gantry crane as the one reported in Fig. 17.17(a). Considering only the advancement direction of the gantry, the schematic representation of this planar system corresponds to the one of a cart-pendulum system (see Fig. 17.17(b)).*

The equations governing the dynamics of this system in the neighbourhood of the downward rest position $\theta \approx 0$ are given in the following

$$\begin{aligned} \frac{d}{dt}v(t) &= -\frac{mL}{M+m}\frac{d}{dt}\omega(t) + \frac{1}{M+m}f(t) \\ \frac{d}{dt}p(t) &= v(t) \\ \frac{d}{dt}\theta(t) &= \omega(t) \\ \frac{d}{dt}\omega(t) &= -\frac{1}{L}\frac{d}{dt}v(t) - \frac{g}{L}\theta(t) \end{aligned} \tag{17.6}$$

The relationship between the force $f(t)$ and the velocity of the gantry $v(t)$ in the Laplace domain is the following one

$$V(s) = \frac{1}{(M+m)s}\frac{s^2 + \dfrac{g}{L}}{s^2 + \dfrac{(M+m)g}{ML}}F(s)$$

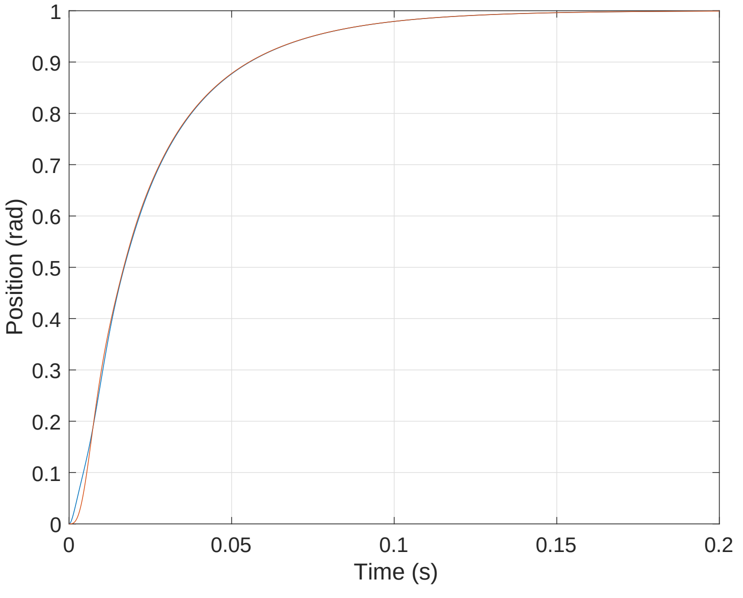
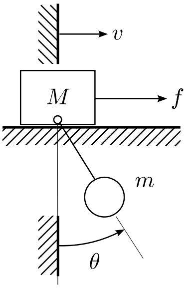


Figure 17.16: Closed-loop unit step response of the motor position $q_m(t)$ (blue) and corresponding scaled load position $\eta q_l(t)$ (red) for the optimal tuning of the velocity controller and the proposed tuning of the position controller.

(a) A heavy-duty gantry crane (Creative Commons)



(b) Schematic model of a planar gantry crane

Figure 17.17: A heavy-duty gantry crane and its planar schematic model.

Notice that the low frequency behaviour corresponds to

$$V(s) \approx \frac{1}{(M+s)\,s} F(s)$$

which corresponds to the load being rigidly attached to the cart, while the transfer function has two undamped complex and conjugate zeros in

$$\omega_z = \sqrt{\frac{g}{L}}$$

and two undamped complex and conjugate poles corresponding to

$$\omega_p = \omega_z \sqrt{1 + \frac{m}{M}} > \omega_z$$

which is completely aligned with the model of the elastic coupling between the two inertias in (17.4).

Assuming the following parameters for the model at hands $M = 16670$ *kg*, $m = 2000$ *kg, and* $L = 5$ *m, the complex and conjugate zeros correspond to frequency* $\omega_z = 1.4007$ *rad/s, while* $\omega_p = 1.4824$ *rad/s, as well as minimal dampings* $\xi_z = 0.05$, $\xi_p = 0.0529$*, the damping of the closed-loop poles as a function of the crossover frequency normalised to the rigid approximation is maximised for* $\tilde{\omega}_{cv} \approx 0.9$ *(see Fig. 17.18). Hence, for the tuning of*

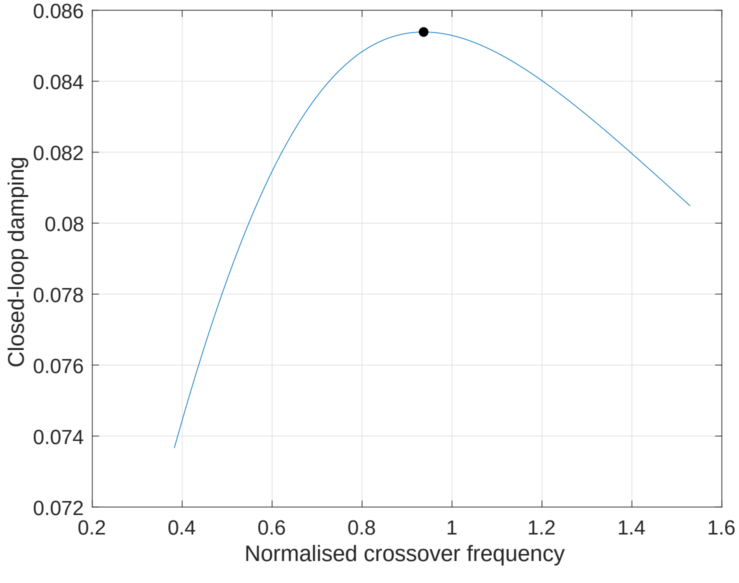


Figure 17.18: Closed-loop damping ξ_v as a function of $\tilde{\omega}_{cv}$.

the velocity controller $R_v(s)$ *we can consider a proportional-integral controller with*

$$T_{iv} = 10/\omega_z = 10.6 \ s, K_{pv} = (M + m)\, 0.9\omega_z = 15868 \ Ns/m$$

Finally, for the tuning of the proportional position controller we can use

$$K_{pp} = 0.09\omega_z = 0.085 \ s^{-1}$$

The step responses of the position of the gantry $p(t)$ *and the corresponding displacement of the load are reported in Fig. 17.19. As somehow expected, the position of the hoisted load does not significantly oscillate with respect to the one of the gantry.*

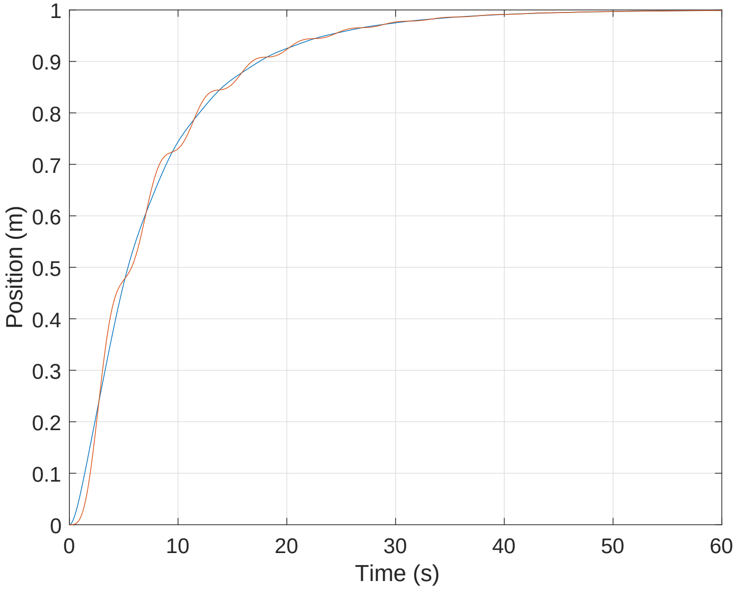


Figure 17.19: Closed-loop unit step response of the gantry position $p(t)$ (blue) and corresponding displacement of the hoisted load (red).

Chapter 18

Motion planning and velocity estimation

18.1 Motion planning

In the previous Chapters, we have seen how to tune the controllers for a mechanical system possibly characterised by some resonant behaviours. All the simulations to validate the effectiveness of the proposed control strategy were referred to step references.

Step references are never used in practice for at least two reasons. First, in the case of an electro-mechanical system the initial value of the required motor voltage will be

$$v\left(0\right) = K_{pi}K_{pv}K_{pp}Q_m^0$$

where Q_m^0 is the amplitude of the position step reference. Depending on the required displacement Q_m^0, the corresponding voltage supply will easily saturate. Similarly, for the case of electric-hydraulic actuations, the initial value of the position of the spool would be

$$x_s^0\left(0\right) = K_p x^0$$

Thought saturations do not represent a particular problem, the presence of a possible oscillating behaviour has to be carefully accounted for.

In Chapter 7, we have seen that the relationship between the reference input, say for example the rotor velocity $\omega_m^0\left(t\right)$ of an electrical motor, and the corresponding output, $\omega_m\left(t\right)$ still referring to an electro-mechanical system, is represented by the complementary sensitivity. Despite any possible effort in maximising the damping of its complex and conjugate poles, this closed-loop transfer function still contains complex and conjugate poles for both the electro-mechanical and the electro-hydraulic systems. It follows that using reference signals characterised by high-frequency components does not certainly represent a good practice. The design of alternative reference signals for the position or the velocity of a mechanical system is therefore the scope of this Section.

18.1.1 Cubic profiles

One of the most intuitive way to generate a reference signal to be used for the outer position control loop is to refer to polynomial functions of time, i.e.

$$q_m^0(t) = a_d t^d + a_{d-1} s t^{d-1} + \cdots + a_1 t + a_0, 0 \leq t \leq t_f$$

Since the number of coefficients of a polynomial of degree d is always $d+1$ and the trajectory has to interpolate quantities at the initial time $t=0$ as well as corresponding quantities at final time $t=t_f$, the order d of polynomials is always an odd number.

For $d=3$, a cubic polynomial has the following form

$$q_m^0(t) = a_3 t^3 + a_2 t^2 + a_1 t + a_0, 0 \leq t \leq t_f$$

and is parametrised in terms of $d+1=4$ coefficients. By imposing initial and final positions, $q_m^0(0), q_m^0(t_f)$ and velocities $\omega_m^0(0), \omega_m^0(t_f)$ as well as the traversal time t_f, one obtains a cubic trajectory characterised by

$$a_0 = q_m^0(0), a_1 = \omega_m^0(0), a_2 = \frac{3q_m(t_f) - \left(2\omega_m^0(0) + \omega_m^0(t_f)\right) t_f}{t_f^2}$$

$$a_3 = \frac{2\left(q_m^0(0) - q_m^0(t_f)\right) + \left(\omega_m^0(0) + \omega_m^0(t_f)\right) t_f}{t_f^3}$$

Example 87. *[Cubic trajectory for the gantry crane] Assume the gantry crane has to move from $p^0(0) = 0\ m$ to $p^0(t_f) = 10\ m$ in $t_f = 60\ s$, starting and arriving with null initial velocities, i.e. with $v^0(0) = v^0(60) = 0\ m/s$, the corresponding cubic position reference can be generated and results*

$$p^0(t) = -9.2593 \cdot 10^{-5} t^3 + 8.3333 \cdot 10^{-3} t^2, 0 \leq t \leq 60$$

As one can see from Fig. 18.1, despite the same positioning time of the trajectory of Fig. 17.19, the motions of both the gantry and the hoisted load appear smoother.

18.1.2 Trapezoidal velocity profiles (TVP)

If the application requires to constrain the maximum velocity and the maximum acceleration along the motion, a candidate solution can be obtained by solving

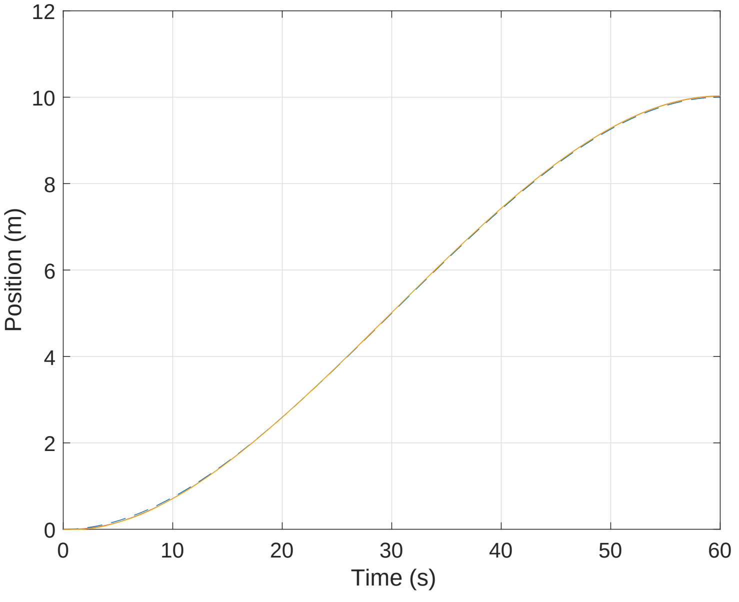


Figure 18.1: Gantry position $p(t)$ (solid red) as compared to its reference $p^0(t)$ (dashed blue) and displacement of the hoisted load (yellow).

the following optimisation problem

$$\begin{aligned} \min_{t_f, q_m^0(t)} t_f &= \min_{t_f, q_m^0(t)} \int_0^{t_f} d\tau \\ \text{subject to } & q_m^0(0) = 0, q_m^0(t_f) \\ & \omega_m^0(0) = 0, \omega_m^0(t_f) = 0 \\ & \left|\omega_m^0(t)\right| \le \omega_m^{\max}, \forall t \in [0, t_f] \\ & \left|\dot{\omega}_m^0(t)\right| \le \dot{\omega}_m^{\max}, \forall t \in [0, t_f] \end{aligned} \tag{18.1}$$

The solution corresponds to a symmetrical trapezoidal velocity profile consisting in three phases: an initial phase with maximal acceleration $\dot{\omega}_m^{\max}$ to reach velocity $\omega_m^{\max}$ having duration

$$t_a = \frac{\omega_m^{\max}}{\dot{\omega}_m^{\max}}$$

a cruised segment with velocity $\omega_m^0(t) = \omega_m^{\max}$, and a final deceleration phase, also with duration t_a, to reach zero final velocity.

Example 88. *[TVP trajectory for the gantry crane] Focusing again on the gantry crane and assuming $v^{\max} = 0.2\ m/s$, and $\dot{v}^{\max} = 0.02\ m/s^2$, the corresponding profile can be computed* *considering*

$$t_a = \frac{v^{max}}{a^{\max}} = 10\ s$$

and

$$p^0\left(t_f\right) = \frac{v^{\max}}{2}\left(t_f + t_f - 2t_a\right)$$

and therefore

$$t_f = \frac{p^0\left(t_f\right)}{v^{\max}} + t_a = 60\ s$$

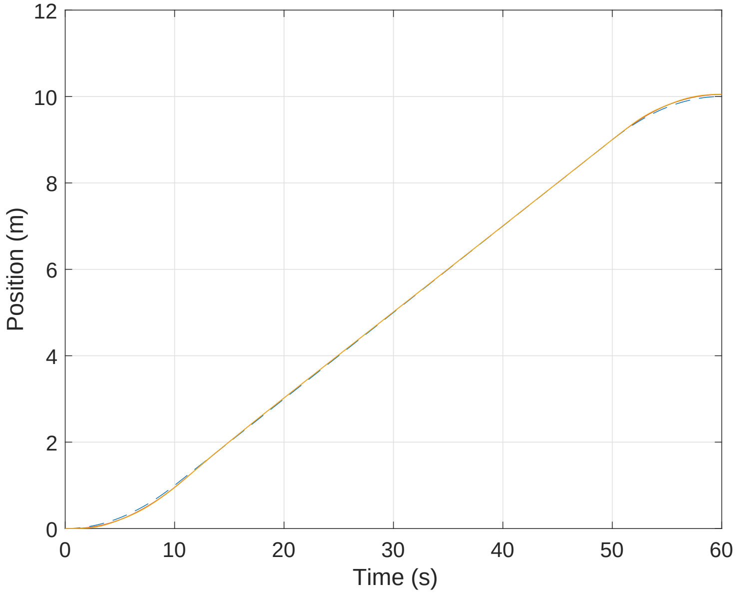


Figure 18.2: Gantry position $p\left(t\right)$ (solid red) as compared to its reference $p^0\left(t\right)$ (dashed blue) and displacement of the hoisted load (yellow).

As one can see from Fig. 18.2, also in this case the motions of both the gantry and the hoisted load appear smoother than the one obtain with a step position reference.

18.1.3 Use of profiles as feed-forward signals

In the previous Chapter, we have introduced the cascaded control architecture for an electro-mechanical system. Recalling the discussion of Chapter 9 on the possibility to complement the feedback architecture with a feed-forward compensator, Fig. 18.3 reports the block-diagram of Fig. 17.4 together with compensator $C_p\left(s\right)$. For brevity, and compactness, the presence of the two internal velocity and current loops are represented by transfer function $F_{vm}\left(s\right)$. The relationship between the Laplace's transform $Q_m^0\left(s\right)$ of the reference position

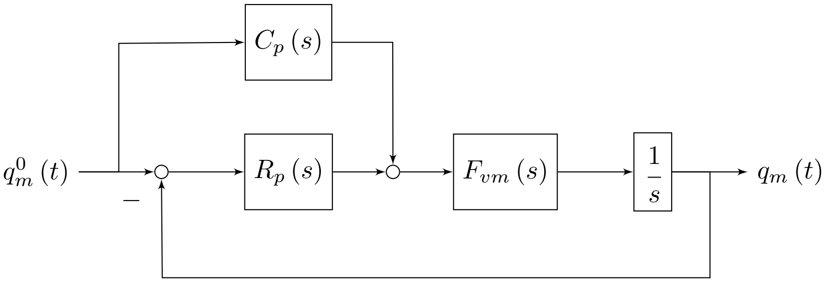


Figure 18.3: Motion control architecture and feed-forward compensator.

and the one of the actual position $Q_m(s)$ can be expressed as

$$Q_m(s) = \frac{R_p(s)\dfrac{F_{vm}(s)}{s} + C_p(s)\dfrac{F_{vm}(s)}{s}}{1 + R_p(s)\dfrac{F_{vm}(s)}{s}} Q_m^0(s)$$

Therefore, the ideal compensator would be the following one

$$C_p(s) = \frac{s}{F_{vm}(s)}$$

Moreover, since $|F_{vm}(j\omega)| \approx 1$ at least for $\omega < \omega_{cp}$ and consistently with the hypothesis behind the adopted cascaded control structure, the ideal compensator can be simplified to

$$C_p(s) = s$$

which is not a transfer function. However, a significant insight can be derived from the ideal compensator. The optimal feed-forward signal for the control scheme in Fig. 18.3 is represented by the derivative of the position reference $q_m^0(t)$. From the discussion on how reference signals are generated, this derivative can be actually calculated in closed-form and results

$$\frac{d}{dt}q_m^0(t) = 3a_3t^2 + 2a_2t + a_1, 0 \le t \le t_f$$

and

$$\frac{d}{dt}q_m^0(t) = \begin{cases} \dot{\omega}_m^{\max} t & t \le t_a \\ \omega_m^{\max} & t_a < t \le t_f - t_a \\ \omega_m^{\max} - \dot{\omega}_m^{\max}(t - t_f + t_a) & t_f - t_a < t \le t_f \end{cases}$$

for the cubic and the trapezoidal velocity profiles, respectively. These values can be directly used as feed-forward terms without the need for approximating and implementing the control code for the transfer function of the feed-forward compensator $C_p(s)$.

18.1.4 Spectral analysis of motion profiles

In the previous examples, we have introduced two alternatives to steer the gantry crane from zero position and velocity, to a displacement of 10 m in 60 s. In the examples, we have seen that in all the cases both the gantry crane and the hoisted load did not present significant oscillations. To better understand which of the candidate motion profiles is capable of reducing the injection of high-frequency components, a spectral analysis of the corresponding acceleration profiles can be set up.

The acceleration profile corresponding to the cubic profile can be written as

$$\dot{\omega}_m^0(t) = 6a_3t + 2a_2, 0 \leq t \leq t_f$$

which can be rewritten as follows

$$\dot{\omega}_m^0(t) = (6a_3t + 2a_2)(t \geq 0) - (6a_3(t - t_f) + 2a_2 - 6a_3t_f)(t \geq t_f) \qquad (18.2)$$

Applying the Laplace's transform to the last polynomial function, one obtains

$$\dot{\Omega}_m^0(s) = \frac{6a_3}{s^2} + \frac{2a_2}{s} - \frac{6a_3e^{-st_f}}{s^2} - \frac{(2a_2 - 6a_3t_f)e^{-st_f}}{s}$$

In the case of the trapezoidal velocity profile, the corresponding acceleration reference can be written as

$$\dot{\omega}_m^0(t) = \dot{\omega}^{\max}((t \geq 0) - (t \geq t_a) - (t \geq t_f - t_a) + (t \geq t_f))$$

which corresponds to the following Laplace's transform

$$\dot{\Omega}_m^0(s) = \frac{\dot{\omega}_m^{\max}}{s} - \frac{\dot{\omega}_m^{\max}e^{-st_a}}{s} - \frac{\dot{\omega}_m^{\max}e^{-s(t_f - t_a)}}{s} + \frac{\dot{\omega}_m^{\max}e^{-st_f}}{s}$$

Given the two Laplace transforms $\dot{\Omega}_m^0(s)$ for the cubic, and the TVP, one can evaluate the spectral content of the corresponding acceleration profiles by simply computing $\left|\dot{\Omega}_m^0(j\omega)\right|$.

Example 89. *[Spectral components of the reference acceleration for the gantry crane] For the cubic and the trapezoidal motion profiles already evaluated for the motion control problem of the gantry crane, the corresponding frequency contents of the reference accelerations $\omega_m^0(t)$ are reported in Fig. 18.4.*

The values of $\left|\dot{\Omega}_m^0(j\omega)\right|$ in the neighbourhood of ω_p are slightly larger in the case of trapezoidal profile. We then expect the trapezoidal velocity profile to introduce frequency components in the neighbourhood of the oscillating modes with sensitively higher amplitudes.

As expected, the trapezoidal velocity profile is responsible for more visible oscillations of the behaviour of velocity of the hoisted load.

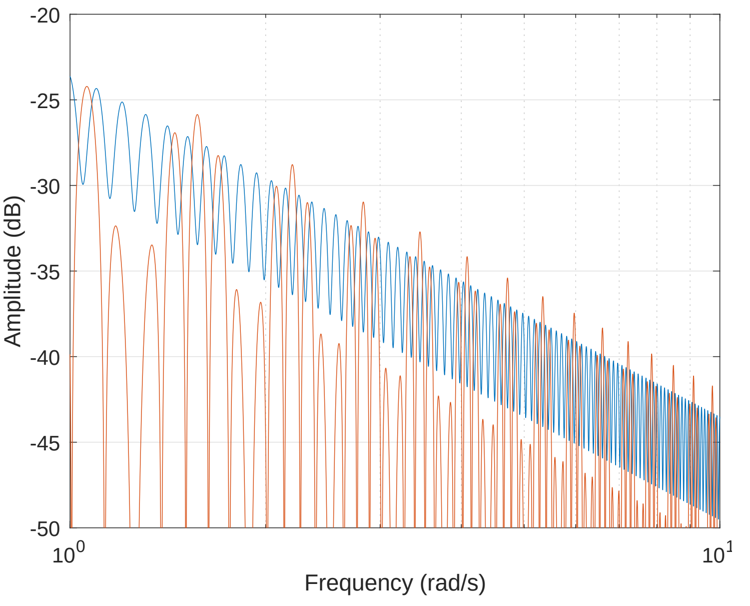


Figure 18.4: Frequency components of the cubic acceleration (blue) and the acceleration of the trapezoidal profile (red).

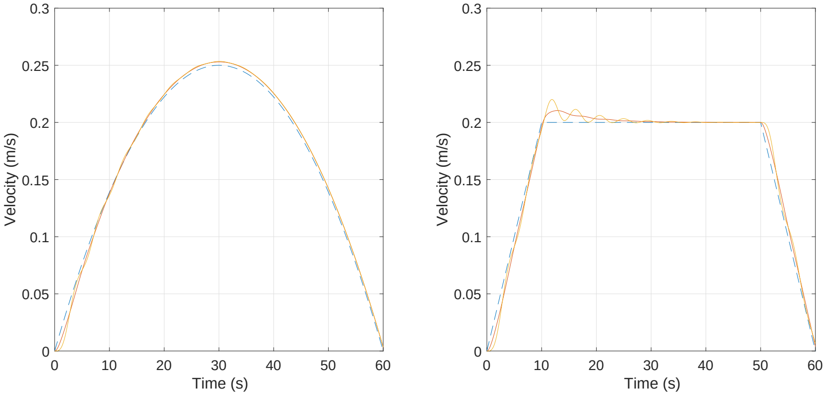


Figure 18.5: Gantry velocity $v(t)$ (solid red) as compared to its reference $v^0(t)$ (dashed blue) and displacement of the hoisted load (yellow) for the cubic (left) and the trapezoidal velocity profile (right).

18.2 Velocity estimation

In the previous Chapter, and in particular in the block diagram of Fig. 17.4, we assumed the velocity of the rotor $\omega_m(t)$ to be directly measurable. In most of the mechatronics applications, however, the velocity feedback is rather based on

an estimate $\hat{\omega}_m(t)$ of the actual velocity of the rotor $\omega_m(t)$. Such an estimate is typically obtained by numerical differentiation of position measurements $q_m(t)$, and sometimes using model-based techniques such as Luenberger observers (or Kalman filters).
Many technologies are available to measure the (angular) position of a rotating shafts: rotary potentiometers, hall sensors, resolver, optimal encoders, etc. No matter which technology is used to acquire the position measurement $q_m(t)$, the corresponding value available within the motion controller will be surely quantized, hence represented with a finite, yet small, resolution, as exemplified in Fig. 18.6.

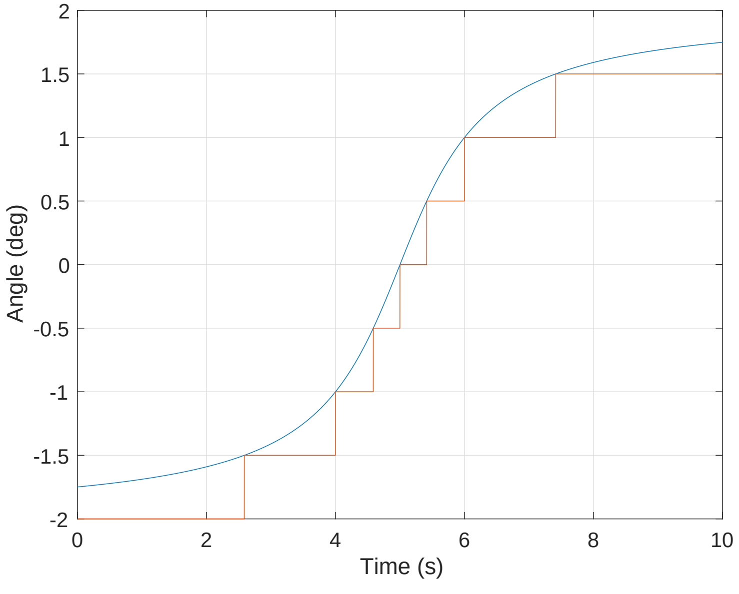


Figure 18.6: Example of motor position $q_m(t)$ (blue) and the corresponding signal after quantisation (red).

It should not be surprising that the value of the position measurement within the controller q_m changes corresponding to discrete events and according to the following discrete-time LTI dynamics

$$q_m(k) = q_m(k-1) + u(k-1)$$

where $u(k-1)$ is equal to the resolution of the sensor, for positive velocities, or its opposite if velocity is negative. Therefore, assuming the two consecutive measures $q_m(k)$ and $q_m(k-1)$ to arrive at time instants $t(k)$ and $t(k-1)$, respectively, an estimate of the velocity $\hat{\omega}_m$ at time instant $t(k)$ can be derived as

$$\hat{\omega}_m\left(t\left(k\right)\right) = \frac{q_m\left(k\right) - q_m\left(k-1\right)}{t\left(k\right) - t\left(k-1\right)}$$

which is based on just two samples $q_m\left(k\right)$ and $q_m\left(k-1\right)$ and the corresponding arriving time instants $t\left(k\right)$ and $t\left(k-1\right)$. More in general, assuming to have $M \geq 2$ measurement samples

$$q_m\left(k\right), q_m\left(k-1\right), \ldots, q_m\left(k-M+1\right)$$

and the corresponding time instants

$$t\left(k\right), t\left(k-1\right), \ldots, t\left(k-M+1\right)$$

it is possible to setup an approximation problem considering a polynomial of order D, with $1 \leq D \leq M-1$, which interpolates or approximates the given samples corresponding to the given time instants. Since a generic polynomial of order D has $D+1$ coefficients, whenever $M = D+1$ the number of samples the problem is an interpolation one, in turn whenever $M < D+1$ the problem corresponds to a polynomial approximation.

A generic polynomial of order D to be used as an approximation of $q_m\left(t\right)$ can be written as follows

$$\hat{q}_m\left(t\right) = x_D t^D + x_{D-1} t^{D-1} + \cdots + x_1 t + x_0 = \begin{bmatrix} t^D & t^{D-1} & \ldots & t & 1 \end{bmatrix} \begin{bmatrix} x_D \\ x_{D-1} \\ \vdots \\ x_1 \\ x_0 \end{bmatrix}$$

then introducing

$$\boldsymbol{x} = \begin{bmatrix} x_D & x_{D-1} & \ldots & x_1 & x_0 \end{bmatrix}^T$$

the vector containing the coefficients of the polynomial and similarly the vector containing time instants and their powers

$$\boldsymbol{a}\left(t\right)^T = \begin{bmatrix} t^D & t^{D-1} & \ldots & t & 1 \end{bmatrix}$$

then, the set $\boldsymbol{x}^*$ of coefficients identifying the interpolating polynomial can be obtained as

$$\boldsymbol{x}^* = \operatorname{argmin}_{\boldsymbol{x}} \sum_{i=0}^{M-1} \left(\hat{q}_m\left(t\left(k-i\right)\right) - q_m\left(k-i\right)\right)^2$$

Moreover, since $\hat{q}_m\left(t\right) = \boldsymbol{a}\left(t\right)^T \boldsymbol{x}$, we have

$$\boldsymbol{x}^* = \operatorname{argmin}_{\boldsymbol{x}} \sum_{i=0}^{M-1} \left(\boldsymbol{a}^T\left(t\left(k-i\right)\right)\boldsymbol{x} - q_m\left(k-i\right)\right)^2 = \operatorname{argmin}_{\boldsymbol{x}} \left\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\right\|^2 \tag{18.3}$$

where

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}\left(t\left(k\right)\right)^T \\ \boldsymbol{a}\left(t\left(k-1\right)\right)^T \\ \vdots \\ \boldsymbol{a}\left(t\left(k-M+1\right)\right)^T \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} q_m\left(k\right) \\ q_m\left(k-1\right) \\ \vdots \\ q_m\left(k-M+1\right) \end{bmatrix}$$

The least squares problem in (18.3) can be finally solved resorting to the Moore-Penrose pseudo-inverse of matrix $\boldsymbol{A}$, i.e.

$$\boldsymbol{x}^* = \boldsymbol{A}^\dagger \boldsymbol{b} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b} \tag{18.4}$$

In fact

$$\left\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\right\|^2 = \left(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\right)^T\left(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\right) = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{b}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{b}$$

and the corresponding gradient with respect to $\boldsymbol{x}$ results

$$\frac{\partial}{\partial \boldsymbol{x}}\left\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\right\|^2 = 2\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}^T\boldsymbol{b}$$

Finally, given the approximating polynomial $\hat{q}_m\left(t\right) = \boldsymbol{a}\left(t\right)^T\boldsymbol{x}$, the corresponding estimate of the velocity at any time t can be computed as

$$\hat{\omega}_m\left(t\right) = \frac{d}{dt}\boldsymbol{a}\left(t\right)^T\boldsymbol{x}$$

The optimal solution in (18.4) can be then obtained by nullifying the gradient.

Example 90. *[Estimation of velocity from quantised position measurements] Consider the following $M = 4$ samples*

$$q_m\left(k\right) = -1 \text{ deg}, q_m\left(k-1\right) = -0.75, q_m\left(k-2\right) = -0.5, q_m\left(k-3\right) = -0.25$$

and the corresponding time instants

$$t\left(k\right) = 7.1068\ s, t\left(k-1\right) = 6.7107, t\left(k-2\right) = 6.2494, t\left(k-3\right) = 5.6569$$

and assume to approximate their values with a polynomial of order $D = 2$, then the corresponding coefficients can be computed as in (18.4), where

$$\boldsymbol{A} = \begin{bmatrix} 50.507 & 7.1068 & 1 \\ 45.033 & 6.7107 & 1 \\ 39.055 & 6.2494 & 1 \\ 32.000 & 5.6569 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} -0.25 \\ -0.5 \\ -0.70 \\ -1 \end{bmatrix}$$

The corresponding polynomial is then the following one:

$$\hat{q}_m\left(t\right) = 0.11012t^2 - 0.88794t + 0.49905$$

while its derivative, i.e. the estimate of the velocity at a generic time instant t, is as follows

$$\hat{\omega}_m\left(t\right) = 0.22023t - 0.88794$$

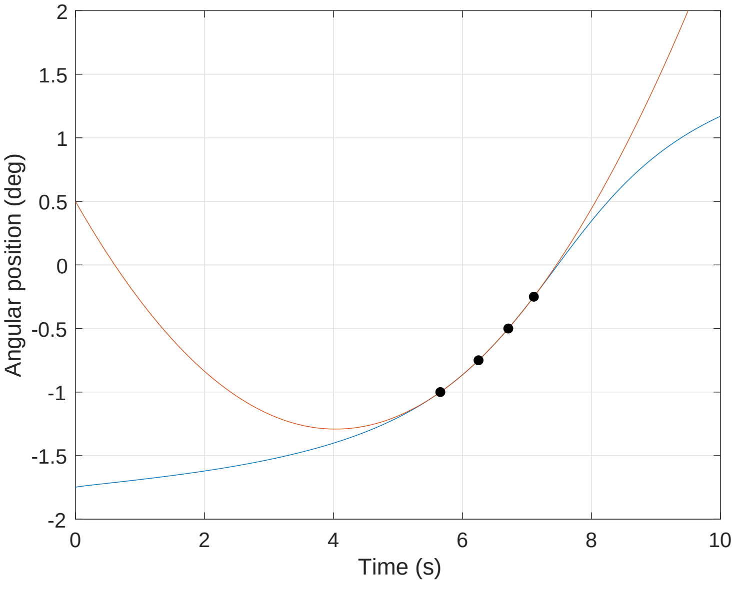


Figure 18.7: Motor position $q_m(t)$ (blue), corresponding samples (black dots) and its approximation using a polynomial of order $D = 2$ (red).

18.3 A complete example

In the following, we are going to present a complete example for the design of a motion control system in the case of electro-mechanical actuators which are typically more demanding in terms of performance with respect to their electro-hydraulic counterpart.

Example 91. *Consider the mechanical system in Fig. 17.2 where with a C23-L33-10 PMDC motor characterised by $k_t = 0.0187\ Nm/A$, and $J_m = 0.0016\ kg \cdot m^2$ connected to a load $J_l = 15.54\ kg \cdot m^2$ with a transmission with ratio $\eta = 100$ and $\omega_z = 500\ rad/s$. The position of the main shaft of the motor is acquired through a $N = 12$ bits encoder.*

We want to design the overall control system that allows the machine to perform a cycle characterised by the following steps which initiate upon the pressure of an enabling button:

1. *acceleration to reach velocity $\omega_m(t) = 10\ rad/s$ after 1 second;*
2. *the velocity of $\omega_m(t) = 10\ rad/s$ is maintained constant for 1 second;*
3. *deceleration to zero velocity of in 1 second;*
4. *standing still for 1 second;*
5. *repeating the previous steps in the opposite direction, i.e. with a target velocity of $\omega_m(t) = -10\ rad/s$.*

The machine then remains at rest until the enabling button is pressed again.

We have already seen in previous examples that the optimal tuning of the proportional-integral velocity controller can be obtained setting

$$T_{iv} = 10/\omega_z = 12.5\ ms, K_{pv} = \left(J_m + \frac{J_l}{\eta^2}\right)\frac{0.8\omega_z}{k_t} = 66.5\ As/rad$$

hence the transfer function of the controller results

$$R_v\left(s\right) = \frac{1.33s + 66.48}{0.02s}$$

As the crossover frequency can be estimated to be

$$\omega_{cv} \approx 0.8\omega_z = 400\ rad/s$$

the sampling time T_s *for the digital implementation of the controller can be selected as*

$$T_s = \frac{\pi}{40\omega_{cv}} = 0.2\ ms$$

This way, the corresponding decrease in the phase margin can be evaluated as

$$\Delta\psi_m = \frac{180}{\pi}\omega_{cv}\left(T_s + \frac{T_s}{2}\right) = 6.75\ \deg$$

which also accounts for the delay to handle the non-strictly-proper controller.

The corresponding discrete-time transfer function can be computed and results

$$R_{vd}\left(z\right) = \frac{66.81z - 66.15}{z-1} = \frac{66.81 - 66.15z^{-1}}{1-z^{-1}}$$

and therefore

$$u\left(k\right) = u\left(k-1\right) + 66.81e\left(k\right) - 66.15e\left(k-1\right)$$

which will be implemented directly in the SFC program.

As per the estimation of the motor velocity $\hat{\omega}_m\left(t\right)$*, the* $N = 12$ *bits encoder has a resolution of*

$$\Delta q_m = \frac{360}{2^{12}} = 0.0879\ \deg$$

We can decide to use a velocity estimator that utilises $M = 5$ *samples and an interpolating polynomial of degree* $D = 3$*. The corresponding estimate of the velocity of the motor will have the following form*

$$\hat{\omega}_m\left(t\right) = 3x_3t^2 + 2x_2t + x_1$$

where the coefficients $x_1, x_2,$ *and* x_3 *are obtained, every time a new position measurement arrives, by solving:*

$$\begin{bmatrix} t\left(k\right)^3 & t\left(k\right)^2 & t\left(k\right) & 1 \\ t\left(k-1\right)^3 & t\left(k-1\right)^2 & t\left(k-1\right) & 1 \\ t\left(k-2\right)^3 & t\left(k-2\right)^2 & t\left(k-2\right) & 1 \\ t\left(k-3\right)^3 & t\left(k-3\right)^2 & t\left(k-3\right) & 1 \\ t\left(k-4\right)^3 & t\left(k-4\right)^2 & t\left(k-4\right) & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_2 \\ x_1 \\ x_0 \end{bmatrix} = \begin{bmatrix} q_m\left(k\right) \\ q_m\left(k-1\right) \\ q_m\left(k-2\right) \\ q_m\left(k-3\right) \\ q_m\left(k-4\right) \end{bmatrix} \tag{18.5}$$

The SFC program for the execution of control code is reported in Fig. 18.9. The upper subprogram represents the update of the digital control law of the proportional-integral velocity controller which runs periodically without interruptions. Notice that step $S0$ *is never deactivated, and therefore* $S0.t$ *basically represents time* t*. Values* $x_1, \ldots, x_3$ *are assumed to be already available in memory as the solution of (18.5).*

In turn, the lower part implements the generation of the reference velocity according to the specified trapezoidal velocity profile and the defined cycle. Notice that the program will remain in state $S1$ *as long as the enabling button is not pressed by an operator.*

Assuming the button is pressed after 1 *second, the time histories of the motor* $\omega_m\left(t\right)$ *and the load* $\omega_l\left(t\right)$ *velocities are reported in Fig. 18.8.*

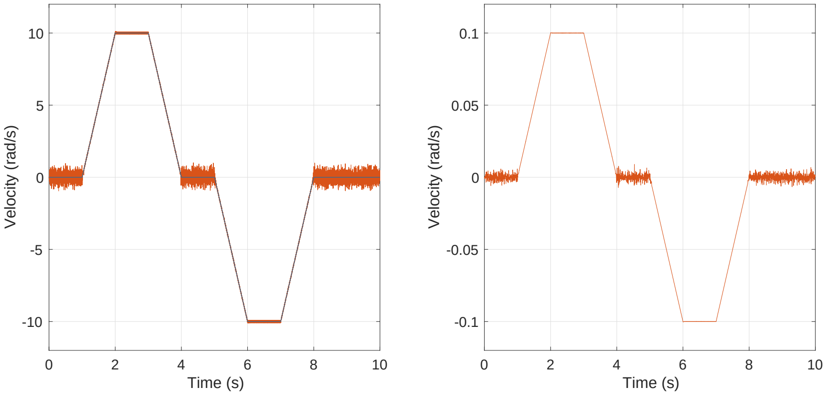


Figure 18.8: Motor velocity $\omega_m(t)$ (red) as compared to its reference $\omega_m^0(t)$ (blue), and corresponding load velocity $\omega_l(t)$ (red, right).

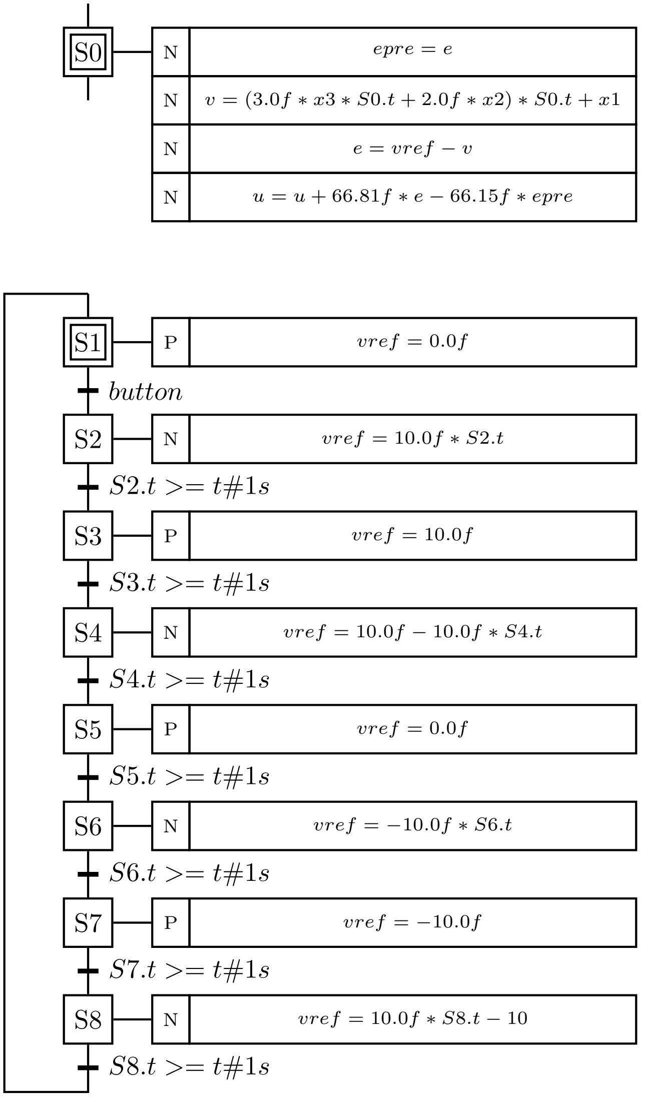


Figure 18.9: SFC code to run the prescribed cycle (executed every $T_s = 0.2$ ms).