

Chapter 3

Ordinary Linear Differential Equations



The automatic control techniques employed in classical control require knowledge of the mathematical model of the physical system to be controlled. As has been shown in Chap. 2, these mathematical models are differential equations. The controller is designed as another differential equation that must be connected in closed-loop to the system to be controlled. This results in another differential equation representing the closed-loop control system. This differential equation is forced by the controller to possess the mathematical properties that ensure that its solution evolves as desired. This means that the controlled variable evolves as desired. Hence, it is important to know the properties of a differential equation determining how its solution behaves in time. Although several different approaches exist to solve differential equations, the use of the Laplace transform is the preferred method in classical control. This is the reason why the Laplace transform method is employed in this chapter, to study linear ordinary differential equations with constant coefficients.

Let $f(t)$ be a function of time, the Laplace transform of $f(t)$ is represented by $F(s)$ and is computed as [2], pp. 185, [3], pp. 285:

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} f(t)e^{-st} dt. \tag{3.1}$$

The fundamental property of the Laplace transform employed to solve linear ordinary differential equations with constant coefficients is the following [2], pp. 192, [3], pp. 293:

$$\mathcal{L}\left\{\frac{d^n f(t)}{dt^n}\right\} = s^n F(s) - s^{n-1} f(0) - s^{n-2} f^{(1)}(0) - \dots - s f^{(n-2)}(0) - f^{(n-1)}(0), \tag{3.2}$$

where the exponent in parenthesis indicates the order of the derivative with respect to time. The following property is also useful [2], pp. 193,[3], pp. 293:

$$\mathcal{L}\left\{\int_0^t f(\tau)d\tau\right\} = \frac{F(s)}{s}. \quad (3.3)$$

Other important properties of the Laplace transform are the final value theorem [1], pp. 25, [3], pp. 304:

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sF(s), \quad (3.4)$$

which is only valid if the indicated limits exist, and the initial value theorem [3], pp. 304:

$$f(0^+) = \lim_{t \rightarrow 0^+} f(t) = \lim_{s \rightarrow \infty} sF(s). \quad (3.5)$$

3.1 First-Order Differential Equation

Consider the following linear, ordinary, first-order, differential equation with constant coefficients:

$$\dot{y} + a y = k u, \quad y(0) = y_0, \quad (3.6)$$

where k and a are real constants, y_0 is a real constant known as the initial condition or initial value of $y(t)$ and $\dot{y} = \frac{dy}{dt}$. The solution of this differential equation is a function of time $y(t)$, which satisfies (3.6) when replaced. With this aim, it is required that constants a , k , y_0 , and the function of time u be known. Use of (3.2) in (3.6) yields:

$$sY(s) - y_0 + a Y(s) = k U(s). \quad (3.7)$$

Solving for $Y(s)$:

$$Y(s) = \frac{k}{s+a}U(s) + \frac{1}{s+a}y_0. \quad (3.8)$$

As stated above, to continue, it is necessary to know u as a function of time. Assume that:

$$u = A, \quad U(s) = \frac{A}{s}, \quad (3.9)$$

where A is a constant. Hence, (3.8) can be written as:

$$Y(s) = \frac{kA}{s(s+a)} + \frac{1}{s+a}y_0. \quad (3.10)$$

According to the partial fraction expansion method [2], Ch. 4, [3], pp. 319, if $a \neq 0$, then it is possible to write:

$$\frac{kA}{s(s+a)} = \frac{B}{s+a} + \frac{C}{s}, \quad (3.11)$$

where B and C are constants to be computed. The value of B can be computed, multiplying both sides of (3.11) by the factor $(s+a)$ and then evaluating the resulting expressions at $s = -a$:

$$\left. \frac{kA}{s(s+a)}(s+a) \right|_{s=-a} = \left. \frac{B}{s+a}(s+a) \right|_{s=-a} + \left. \frac{C}{s}(s+a) \right|_{s=-a},$$

i.e.,

$$B = \left. \frac{kA}{s} \right|_{s=-a} = -\frac{kA}{a}. \quad (3.12)$$

The value of C is computed multiplying both sides of (3.11) by the factor s and then evaluating the resulting expressions at $s = 0$:

$$\left. \frac{kA}{s(s+a)}s \right|_{s=0} = \left. \frac{B}{s+a}s \right|_{s=0} + \left. \frac{C}{s}s \right|_{s=0},$$

i.e.,

$$C = \left. \frac{kA}{s+a} \right|_{s=0} = \frac{kA}{a}. \quad (3.13)$$

Replacing (3.11), (3.12), (3.13) in (3.10) yields:

$$Y(s) = -\frac{kA}{a} \frac{1}{s+a} + \frac{kA}{a} \frac{1}{s} + \frac{y_0}{s+a}. \quad (3.14)$$

From tabulated Laplace transform formulas [4], Ch. 32, it is known that:

$$\mathcal{L}\{e^{\beta t}\} = \frac{1}{s-\beta}, \quad (3.15)$$

where β is any real constant. Using (3.15) it is found that the inverse Laplace transform of (3.14) is given as:

$$y(t) = -\frac{kA}{a}e^{-at} + \frac{kA}{a} + y_0e^{-at}, \quad (3.16)$$

which is the solution of (3.6).

Example 3.1 To verify that (3.16) is the solution of (3.6) proceed as follows. First, compute the first time derivative of $y(t) = -\frac{kA}{a}e^{-at} + \frac{kA}{a} + y_0e^{-at}$, i.e.,

$$\dot{y}(t) = kAe^{-at} - ay_0e^{-at},$$

and replace $\dot{y}(t)$ and $y(t)$ in (3.6) to obtain:

$$\dot{y}(t) + ay(t) = kAe^{-at} - ay_0e^{-at} + a\left(-\frac{kA}{a}e^{-at} + \frac{kA}{a} + y_0e^{-at}\right) = kA = ku,$$

because $u = A$ has been defined in (3.9). This means that the expression for $y(t)$ given in (3.16) satisfies (3.6); thus, it has been verified that (3.16) is the solution of (3.6). This procedure must be followed to verify the solution of any differential equation.

The solution given in (3.16) can be decomposed into two parts:

$$y(t) = y_n(t) + y_f(t), \quad (3.17)$$

$$y_n(t) = \left(-\frac{kA}{a} + y_0\right)e^{-at}, \quad (3.18)$$

$$y_f(t) = \frac{kA}{a}, \quad (3.19)$$

where $y_n(t)$ and $y_f(t)$ are called *the natural response* and *the forced response* respectively. The forced response $y_f(t)$ receives this name because it is caused by the presence of the function of time u . On the other hand, the natural response $y_n(t)$ receives this name because it is produced by the structure (i.e., the nature) of the differential equation, i.e., by terms $\dot{y} + ay$ on the left hand of (3.6). Let us explain this in more detail. Follow the solution procedure presented above, but assume for a moment that $y_0 = 0$. The partial fraction expansion presented in (3.11) and the subsequent application of the inverse Laplace transform show that $y(t)$ is given as the sum of several functions of time that result from the fractions $\frac{1}{s}$ and $\frac{1}{s+a}$. Notice that the presence of $\frac{1}{s}$ in (3.11) is due to the fact that $U(s) = \frac{A}{s}$ appears in (3.10), which is because $u = A$ where A is a constant. Notice that the forced response $y_f(t)$ is a constant resulting from the term $\frac{kA}{a} \frac{1}{s}$ in (3.14), i.e., it arises from fraction $\frac{1}{s}$. This corroborates that $y_f(t)$ only depends on the function of time u .

On the other hand, fraction $\frac{1}{s+a}$ is due to terms $\dot{y} + ay$ as $\mathcal{L}\{\dot{y} + ay\} = (s+a)Y(s)$. According to (3.15), the consequence of this fraction is function e^{-at} , which constitutes the natural response $y_n(t)$. This corroborates that $y_n(t)$ is only

determined by the differential equation structure (i.e., its nature). One important consequence of this fact is that $y_n(t)$ is always given by the function e^{-at} no matter what the value of u is. The reader can review the procedure presented above to solve this differential equation, thereby realizing that the initial condition y_0 always appears as a part of the natural response $y_n(t)$.

In control systems, y and u are known as the output and the input respectively. The polynomial resulting from the application of the Laplace transform in the differential equation, i.e., $s + a$, which arises from $\mathcal{L}\{\dot{y} + ay\} = (s + a)Y(s)$, is known as the characteristic polynomial. The solution in (3.16), or equivalently in (3.17), can evolve according to one of the following cases.

1. If $a > 0$, i.e., if the only root of the characteristic polynomial $s = -a$ is real and negative, then $\lim_{t \rightarrow \infty} y_n(t) = 0$ and $\lim_{t \rightarrow \infty} y(t) = y_f(t)$.
2. If $a < 0$, i.e., if the only root of the characteristic polynomial $s = -a$ is positive, then $y_n(t)$ and $y(t)$ grow without limit as time increases.
3. The case when $a = 0$, i.e., when the only root of the characteristic polynomial is zero $s = 0$, cannot be studied from results that have been obtained so far and it is studied as a special case in the next section. However, to include all the cases, let us talk about what is to be found in the next section. When the only root of the characteristic polynomial is at $s = -a = 0$, the natural response is constant $y_n(t) = y_0, \forall t \geq 0$ and the forced response is the integral of the input $y_f(t) = k \int_0^t u(t) dt$.

3.1.1 Graphical Study of the Solution

According to the previous section, if $a > 0$, then the natural response goes to zero over time: $\lim_{t \rightarrow \infty} y_n(t) = 0$; hence, when time is large enough the solution of the differential equation is equal to the forced response: $\lim_{t \rightarrow \infty} y(t) = y_f(t) = \frac{kA}{a}$. This means that the faster the natural response $y_n(t)$ tends toward zero (which occurs as $a > 0$ is larger), the faster the complete solution $y(t)$ reaches the forced response $y_f(t)$. Hence, the natural response can be seen as a means of transport, allowing the complete solution $y(t)$ to go from the initial value $y(0) = y_0$ to the forced response $y_f(t)$. Existence of such a means of transport is justified recalling that the solution of the differential equation in (3.6), i.e., $y(t)$, is a continuous function of time. This means that the solution cannot go from $y(t) = y_0$ to $y(t) = y_f(t) \neq y_0$ in a zero time interval.

The graphical representation of the solution in (3.17), when $y_0 = 0$ and $a > 0$, i.e.,

$$y(t) = \frac{kA}{a} (1 - e^{-at}), \quad (3.20)$$

is depicted in Fig. 3.1. An important parameter in first-order differential equations is *time constant* τ , which is defined as:

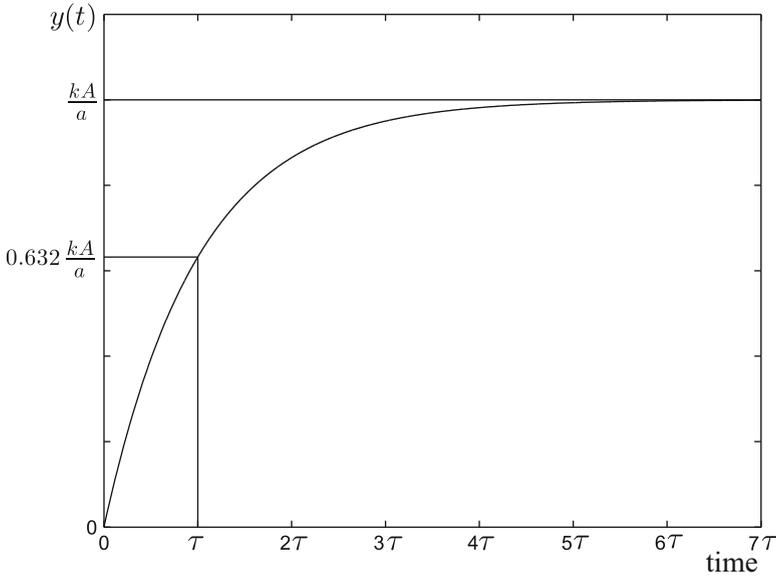


Fig. 3.1 Graphical representation of $y(t)$ in (3.20)

$$\tau = \frac{1}{a}. \quad (3.21)$$

This is indicative of how fast the natural response vanishes, i.e., how fast $y(t)$ approaches the forced response $y_f(t)$. Replacing (3.21) into (3.20), it is easy to verify that:

$$y(\tau) = 0.632 \frac{kA}{a}. \quad (3.22)$$

Finally, it is important to stress that $y(t)$ grows without a limit if $a < 0$. If $a = 0$, then $y(t) = kAt$ also grows without a limit (see Sect. 3.2).

3.1.2 Transfer Function

Consider a zero initial condition, i.e., $y_0 = 0$, then (3.8) can be written as:

$$Y(s) = \frac{k}{s+a} U(s),$$

or, equivalently:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{k}{s+a}. \quad (3.23)$$

The function $G(s)$ is known as the *transfer function* of the differential equation in (3.6). The polynomial in the denominator of $G(s)$, i.e., $s + a$, is known as the characteristic polynomial and its roots are known as the poles of $G(s)$. In this case, there is only one pole at $s = -a$. According to the discussion above, the following can be stated with respect to $y(t)$, i.e., the output of the transfer function $G(s)$:

- If pole located at $s = -a$ is negative, i.e., if $a > 0$, then the natural response $y_n(t)$ vanishes and the complete solution $y(t)$ approaches the forced response $y_f(t)$ as time increases. If $u(t) = A$ is a constant, then $y_f(t) = \frac{kA}{a}$.
- The faster $y(t)$ approaches $y_f(t)$, the farther to the left of the origin $s = 0$ is placed the pole at $s = -a$. This can be quantified using the time constant $\tau = \frac{1}{a}$. A large time constant implies a slow response whereas a small time constant implies a fast response.
- If $k = a > 0$, then it is said that the transfer function $G(s)$ has a unitary gain in a steady state because, according to the final value theorem (3.4), if $u(t) = A$ is a constant then:

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} sY(s) = \lim_{s \rightarrow 0} s \frac{k}{s+a} \frac{A}{s} = \frac{k}{a} A = A. \quad (3.24)$$

In general, the steady-state gain of the transfer function in (3.23) is computed as $\frac{k}{a}$.

- If the pole located at $s = -a$ is positive, i.e., if $a < 0$, then the complete solution $y(t)$ grows without limit. Notice that this implies that the pole is located on the right half of the plane s (see Fig. 3.2).
- The case when the pole is located at $s = a = 0$ cannot be studied from the above computations and it is analyzed as a special case in the next section. However, again to summarize all the possible cases, we present here results that are obtained in the next section. When the pole is located at $s = a = 0$, the natural response neither vanishes nor increases without limit and the forced response is given as the integral of the input $y_f(t) = k \int_0^t u(t) dt$.

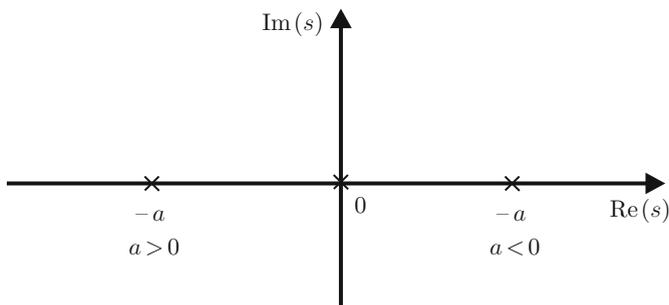
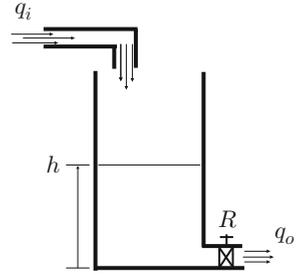


Fig. 3.2 Location of the poles of the transfer function in (3.23). It is usual to represent a pole on the s plane using a cross “ \times ”

Fig. 3.3 A water level system



Example 3.2 Consider the tank containing water depicted in Fig. 3.3. Assume that the tank section C is constant. Water enters the tank at a rate given by the input flow q_i (m^3/s). Water leaves the tank at a rate given by the output flow q_o (m^3/s) through a valve with hydraulic resistance R . The water level in the tank is represented by h . The mathematical model describing this system is obtained using the mass conservation law. As water is not compressible, this law can be stated in terms of mass or volume.

Let ΔV_i and ΔV_o be the water volumes entering and leaving the tank respectively during a time interval Δt . During this time interval, the water volume ΔV increases inside the tank at a rate given as:

$$\frac{\Delta V}{\Delta t} = \frac{\Delta V_i}{\Delta t} - \frac{\Delta V_o}{\Delta t}. \quad (3.25)$$

On the other hand:

$$\Delta V = C \Delta h,$$

$$\Delta V_i = q_i \Delta t,$$

$$\Delta V_o = q_o \Delta t.$$

Replacing this in (3.25) and considering small increments of time such that $\Delta t \rightarrow dt$ and $\Delta h \rightarrow dh$:

$$C \frac{dh}{dt} = q_i - q_o. \quad (3.26)$$

Assuming that the water flow is laminar through the output valve [1], pp. 155, then:

$$q_o = \frac{h}{R}. \quad (3.27)$$

To understand the reason for this expression, consider the following situations.

- a) Suppose that the opening of the output valve is kept without change, i.e., R remains constant, and the water level h increases. Everyday experience and the expression in (3.27) corroborate that the output flow q_o increases under this situation.
- b) Suppose that the water level h is kept constant, because there is water entering the tank to exactly compensate for water leaving the tank. Then, slowly close the output valve, i.e., slowly increase R . Everyday experience and the expression in (3.27) corroborate that the output flow q_o decreases in this situation.

Substituting (3.27) in (3.26), it is found that:

$$C \frac{dh}{dt} = q_i - \frac{h}{R}.$$

Hence, the mathematical model of the water level system depicted in Fig. 3.3 is given as:

$$C \frac{dh}{dt} + \frac{1}{R}h = q_i.$$

Dividing by C , the following linear, ordinary, first-order differential equation with constant coefficients is finally obtained:

$$\begin{aligned} \frac{dh}{dt} + ah &= kq_i, \\ a &= \frac{1}{RC}, \quad k = \frac{1}{C}. \end{aligned} \quad (3.28)$$

The reason why this differential equation is the mathematical model of the water level system is because, solving it, the evolution over time of the water level $h(t)$ can be known if the input flow $q_i(t)$, the tank section C , the hydraulic resistance R , and the initial water level are known.

Example 3.3 Consider the water level system depicted in Fig. 3.3. It was found in Example 3.2 that the mathematical model is given as:

$$\begin{aligned} \frac{dh}{dt} + ah &= kq_i, \\ a &= \frac{1}{RC} > 0, \quad k = \frac{1}{C}. \end{aligned} \quad (3.29)$$

This differential equation can be written as in (3.6) if it is assumed that $y = h$, $u = q_i$ and $y_0 = h_0$. Hence, if $q_i = A$, the solution $h(t)$ is similar to that in (3.16), i.e.,:

$$h(t) = -\frac{kA}{a}e^{-at} + \frac{kA}{a} + h_0e^{-at}, \quad h_0 = h(0). \quad (3.30)$$

Let us analyze (3.30).

1. Assume that the tank is initially empty, i.e., $h_0 = 0$, and water enters the tank with a constant rate $q_i = A > 0$. Under these conditions, the water level $h(t)$ evolves as depicted in Fig. 3.1 by assuming that $h(t) = y(t)$. Notice that $a = \frac{1}{RC} > 0$. From this figure, the solution in (3.30), and using everyday experience, several cases can be studied.
 - If A is chosen to be larger, then the final water level $\lim_{t \rightarrow \infty} h(t) = \frac{kA}{a} = RA$ is also larger.
 - If the output valve is slowly closed, i.e., R is slowly increased, then the final water level $\lim_{t \rightarrow \infty} h(t) = \frac{kA}{a} = RA$ slowly increases.
 - The tank cross-section has no effect on the final water level, i.e., the final water level is the same in a thin tank and a thick tank.
 - If the product RC is larger, then a is smaller; hence, the system response is slower, i.e., more time is required for the level to reach the value $h(\tau) = 0.632 \frac{kA}{a} = 0.632RA$ and, as consequence, also to reach the final water level $\lim_{t \rightarrow \infty} h(t) = \frac{kA}{a} = RA$. This is because the function e^{-at} tends toward zero more slowly when $a = \frac{1}{RC} > 0$ is smaller. Notice that a larger value of RC can be obtained by increasing the tank cross-section C , or decreasing the output valve opening, i.e., increasing R .
2. Suppose that no water is entering the tank, i.e., $q_i(t) = A = 0$, and the initial water level is not zero, i.e., $h_0 > 0$. From the solution in (3.30), and everyday experience can verify it, the water level decreases (recall that $a = \frac{1}{RC} > 0$) until the tank is empty, i.e., $\lim_{t \rightarrow \infty} h(t) = 0$ (see Fig. 3.4). This behavior is faster as RC is smaller, i.e., when a is larger, and slower as RC is larger, i.e., when a is smaller.

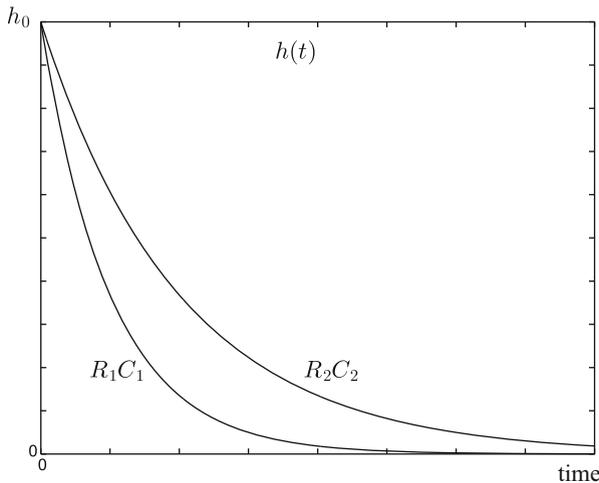


Fig. 3.4 Water level evolution when $h_0 > 0$ and $q_i(t) = A = 0$ ($R_1 C_1 < R_2 C_2$)

3. The case when $a < 0$ is not possible in an open-loop water level system because C and R cannot be negative. However, $a < 0$ may occur in closed-loop systems as a consequence of the interconnection of diverse components.
4. The case when $a = 0$ is studied in the next section.

3.2 An Integrator

Consider the following differential equation:

$$\dot{y} = ku, \quad (3.31)$$

where k is a real constant different from zero. Notice that this equation is obtained as a particular case of (3.6) when $a = 0$. By direct integration:

$$\begin{aligned} \int_{y(0)}^{y(t)} dy &= \int_0^t ku \, dt, \\ y(t) &= k \int_0^t u(t) \, dt + y(0), \\ y(t) &= y_n(t) + y_f(t), \\ y_n(t) &= y(0), \quad y_f(t) = k \int_0^t u(t) \, dt. \end{aligned} \quad (3.32)$$

In this case, $y_n(t)$ remains constant in time. Notice that, if $u = A$ is a constant, the forced response increases without a limit $y_f(t) = kAt$. The solution $y(t) = y(0) + kAt$ is depicted in Fig. 3.5 for a) $y(0) \neq 0$ and $u = A > 0$ a constant, and b) $y(0) \neq 0$ and $u = A = 0$. Physical systems represented by this class of equation are called *integrators* because, if $y(0) = 0$, the solution $y(t)$ is the integral of the excitation u .

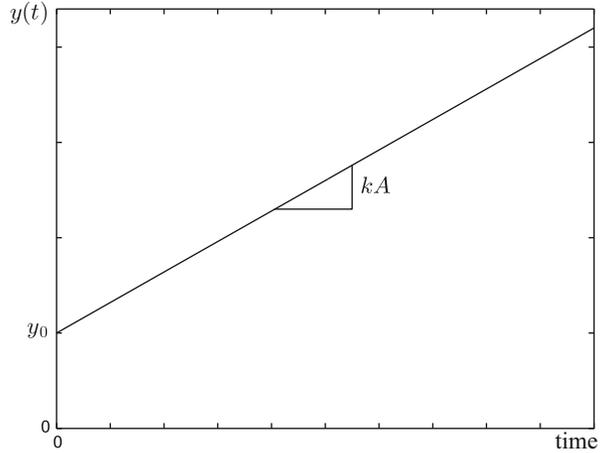
Example 3.4 Consider the water level system depicted in Fig. 3.3. Suppose that the output valve is closed, i.e., $R \rightarrow \infty$. Then, the mathematical model in (3.28) becomes:

$$\frac{dh}{dt} = kq_i,$$

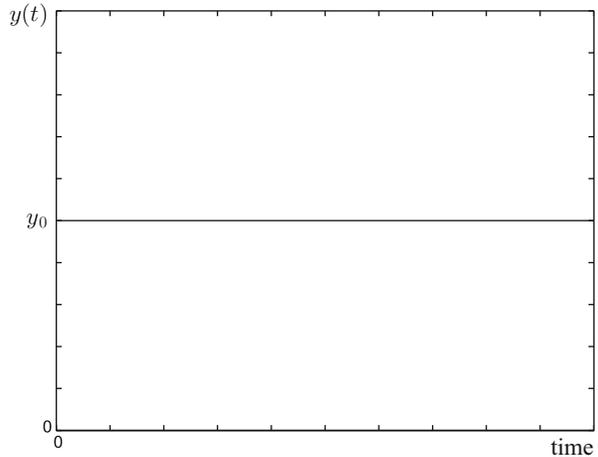
because $a = \frac{1}{RC} \rightarrow 0$ when $R \rightarrow \infty$. Then, the level evolution over time is described by (3.32), i.e.,:

$$h(t) = h(0) + k \int_0^t q_i(t) \, dt.$$

Fig. 3.5 Solution of differential equation in (3.31). (a) $y_0 \neq 0$ and $u = A > 0$ is a constant. (b) $y_0 \neq 0$ and $u = A = 0$



(a)



(b)

Hence, the water level system behaves in this case as an integrator. If $q_i = A > 0$, then:

$$h(t) = h(0) + kAt. \tag{3.33}$$

Figure 3.5 can be employed, assuming $h(t) = y(t)$ and $q_i(t) = u$, to graphically depict the solution in (3.33) for the cases *i*) $h(0) > 0$ and $q_i = A > 0$ is a constant, and *ii*) $h(0) > 0$ and $q_i = A = 0$. Notice that the water level remains constant if no water enters the tank, i.e., when $A = 0$, and the level increases without a limit when $A > 0$ is constant. The reader can resort to everyday experience to verify that these two situations actually occur in a practical water level system.

Example 3.5 The differential equation in (3.31) can also be solved using partial fraction expansion, i.e., using the method employed in Sect. 3.1. From (3.2) and (3.31) it is found that:

$$sY(s) - y_0 = kU(s).$$

Solving for $Y(s)$ yields:

$$Y(s) = \frac{k}{s}U(s) + \frac{1}{s}y_0. \quad (3.34)$$

Assume that:

$$u = A, \quad U(s) = \frac{A}{s},$$

where A is a constant. Hence, (3.34) can now be written as:

$$Y(s) = \frac{kA}{s^2} + \frac{1}{s}y_0. \quad (3.35)$$

According to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, it must be written:

$$\frac{kA}{s^2} = \frac{B}{s} + \frac{C}{s^2}, \quad (3.36)$$

where B and C are two constants. C is computed multiplying both sides of (3.36) by s^2 and evaluating at $s = 0$, i.e.,:

$$s^2 \frac{kA}{s^2} \Big|_{s=0} = \frac{B}{s} s^2 \Big|_{s=0} + \frac{C}{s^2} s^2 \Big|_{s=0}.$$

Hence, $C = kA$ is found. B is computed multiplying both sides of (3.36) by s^2 , differentiating once with respect to s and evaluating at $s = 0$:

$$\frac{d}{ds} \left(s^2 \frac{kA}{s^2} \right) \Big|_{s=0} = \frac{d}{ds} \left(\frac{B}{s} s^2 \right) \Big|_{s=0} + \frac{d}{ds} \left(\frac{C}{s^2} s^2 \right) \Big|_{s=0}.$$

Hence, $B = 0$ is found. Thus, using these values and (3.36), the expression in (3.35) can be written as:

$$Y(s) = \frac{kA}{s^2} + \frac{1}{s}y_0. \quad (3.37)$$

From tabulated formulas, [4], Ch. 32, it is well known that:

$$\mathcal{L}\{t\} = \frac{1}{s^2}. \quad (3.38)$$

Using (3.38), it is found that the inverse Laplace transform of (3.37) is given as:

$$y(t) = kAt + y_0, \quad (3.39)$$

which is the solution of (3.31). This solution can be decomposed into two parts:

$$\begin{aligned} y(t) &= y_n(t) + y_f(t), & (3.40) \\ y_n(t) &= y_0, \\ y_f(t) &= kAt, \end{aligned}$$

where $y_n(t)$ and $y_f(t)$ are the natural and the forced responses respectively. As explained Sect. 3.1, the natural response is given by the Laplace transform of term \dot{y} on the left-hand side of (3.31). As a consequence of this, the term $\frac{1}{s}y_0$ appears in (3.37), i.e., $y_n(t) = y_0$ in (3.40). The forced response $y_f(t)$ is produced by $u = A$. However, in this case, the forced response also receives the effect of the Laplace transform of the term \dot{y} on the left-hand side of (3.31), as the combination of both of them results in $\frac{kA}{s^2}$ and, hence, $y_f(t) = kAt$.

The characteristic polynomial of the differential equation in (3.31) is s ; hence, it only has one root at $s = 0$. Thus, the corresponding transfer function $G(s)$, when $y_0 = 0$:

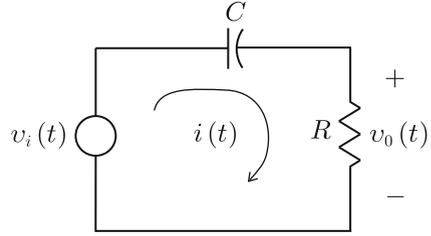
$$\frac{Y(s)}{U(s)} = G(s) = \frac{k}{s}, \quad (3.41)$$

only has one pole at $s = 0$. Notice that the input $U(s) = \frac{A}{s}$ also has one pole at $s = 0$. The combination of these repeated poles in (3.35) produces the forced response $y_f(t) = kAt$, i.e., a first-degree polynomial of time, whereas the input $u = A$ is a zero-degree polynomial of time. Notice that the forced response corresponding to the differential equation in (3.6) is $y_f(t) = \frac{kA}{a}$, i.e., a zero-degree polynomial of time, whereas the input is $u = A$, i.e., a zero-degree polynomial of time. From these examples we arrive at the following important conclusion:

“The forced response of a first-order differential equation with a constant input, is also a constant if its characteristic polynomial has no root at $s = 0$.”

In Fig. 3.2 the location, on the s plane of the pole of the transfer function in (3.41) is depicted, i.e., the pole at $s = 0$. The reader must learn to relate the pole location on the s plane to the corresponding time response $y(t)$. See the cases listed before Example 3.2.

Example 3.6 Consider the electric circuit depicted in Fig. 3.6. According to the Example 2.9, the mathematical model of this circuit is given as:

Fig. 3.6 A series RC circuit

$$\dot{v}_0 + \frac{1}{RC}v_0 = \dot{v}_i.$$

Suppose that $v_i(t)$ is a step signal with amplitude A , i.e., $V_i(s) = \frac{A}{s}$. The solution $v_0(t)$ can be found to proceed as follows. Using the Laplace transform in the previous expression:

$$\begin{aligned} V_0(s) \left(s + \frac{1}{RC} \right) &= sV_i(s) - v_i(0) + v_0(0), \\ V_0(s) &= \frac{s}{s + \frac{1}{RC}} V_i(s) + \frac{v_0(0) - v_i(0)}{s + \frac{1}{RC}}. \end{aligned}$$

Replacing $V_i(s) = \frac{A}{s}$ and using the inverse Laplace transform:

$$\begin{aligned} V_0(s) &= \frac{s}{s + \frac{1}{RC}} \frac{A}{s} + \frac{v_0(0) - v_i(0)}{s + \frac{1}{RC}}, \\ V_0(s) &= \frac{A}{s + \frac{1}{RC}} + \frac{v_0(0) - v_i(0)}{s + \frac{1}{RC}}, \\ v_0(t) &= (A + v_0(0) - v_i(0))e^{-\frac{1}{RC}t}. \end{aligned}$$

Applying KVL at $t = 0$, we have:

$$v_i(0) = v_0(0) + v_c(0),$$

where $v_c(0)$ represents the initial voltage at the capacitor. Then, we have:

$$v_0(t) = (A - v_c(0))e^{-\frac{1}{RC}t}. \quad (3.42)$$

In Fig. 3.7 this time response, when $v_i(t)$ is the square wave represented by the dashed line and the time constant is $RC = 0.1$ [s], is represented. At $t = 0$ we have $v_c(0) = 0$ and $A = 1$. Thus:

$$v_0(t) = Ae^{-10t}. \quad (3.43)$$

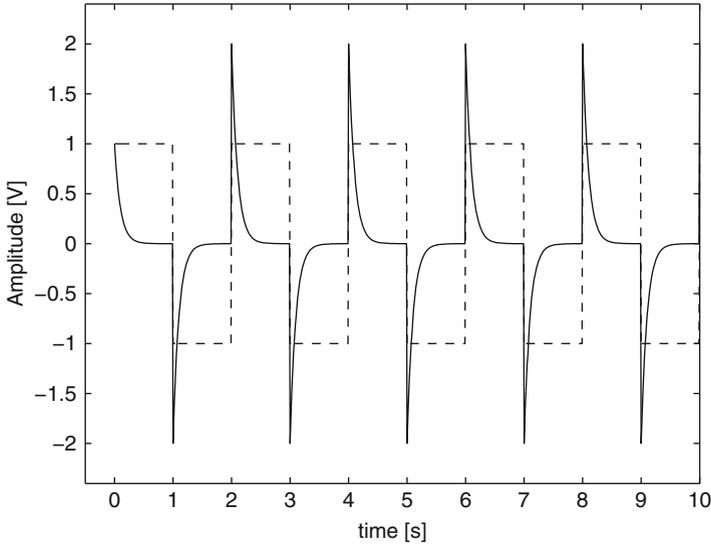


Fig. 3.7 Time response of the circuit in Fig. 3.6. Continuous: $v_0(t)$. Dashed: $v_i(t)$

This response is observed between $t = 0$ and $t = 1$ [s]. Notice that $v_0(0) = A = 1$ [V] is correctly predicted by (3.43). Also notice that $t = 1$ [s] represents ten times the time constant, i.e., $RC = 0.1$ [s]. Hence, $v_0(t)$ is very close to zero at $t = 1$ [s]. Applying the KVL at $t = t_1 = 1$ [s]:

$$\begin{aligned} v_i(t_1) &= v_0(t_1) + v_c(t_1), \\ v_c(t_1) &= A = 1[\text{V}], \end{aligned} \quad (3.44)$$

as $v_i(t_1) = A = 1$ [V] and $v_0(t_1) = 0$.

On the other hand, from $t = t_1 = 1$ [s] to $t = 2$ [s] we have that $A = -1$. Note that (3.42) is still valid in this time interval if $t = t - t_1$ is redefined and (3.42) is rewritten as:

$$\begin{aligned} v_0(t - t_1) &= (A - v_c(t_1))e^{-\frac{1}{RC}(t-t_1)}, \\ v_0(t - t_1) &= -2e^{-10(t-t_1)}, \end{aligned} \quad (3.45)$$

as $v_c(t_1) = 1$ [V] (see (3.44)) and $A = -1$ [V] at $t = t_1$ [s]. In Fig. 3.7, it is observed that $v_0 = -2$ [V], at $t = t_1$, is correctly predicted by (3.45). Also notice that, again, the time interval length 1[s] represents ten times the time constant, i.e., $RC = 0.1$ [s]. Hence, $v_0(t - t_1)$ is very close to zero at $t = 2$ [s]. The above procedure can be repeated to explain the results in Fig. 3.7 for any $0 \leq t \leq 10$.

Figure 3.7 has been obtained using the MATLAB/Simulink diagram presented in Fig. 3.8. The *signal generator block* is programmed to have a square wave form,

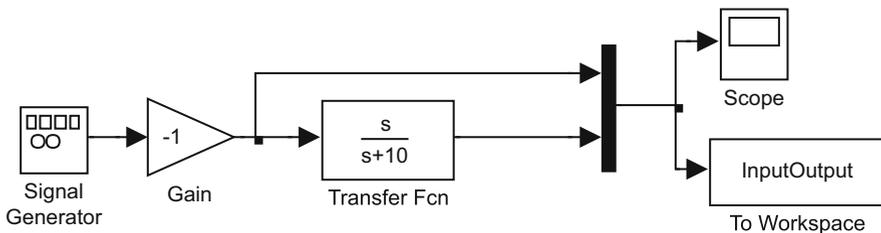


Fig. 3.8 MATLAB/Simulink diagram for the circuit in Fig. 3.6

unit amplitude, and 0.5[Hz] as the frequency. The negative unitary gain block is included because the signal delivered by the signal generator block starts with a negative semicircle. Once the simulation stops, the following MATLAB code is executed in an m-file to draw Fig. 3.7:

```

nn=length(InputOutput(:,1));
n=nn-1;
Ts=10/n;
t=0:Ts:10;
plot(t,InputOutput(:,1),'k--',t,InputOutput(:,2),'k-')
axis([-0.5 10 -2.5 2.5])
    
```

Notice that the transfer function of the circuit in Fig. 3.6 is:

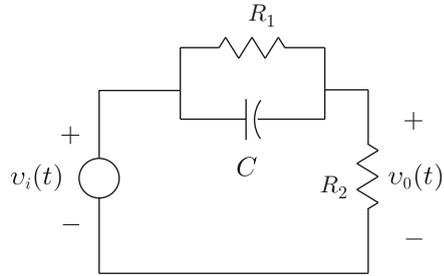
$$\frac{V_0(s)}{V_i(s)} = G(s) = \frac{s}{s + \frac{1}{RC}}, \tag{3.46}$$

when $v_0(0) = 0$ and $v_i(0) = 0$, or equivalently, when $v_c(0) = 0$. Although this is a first-order transfer function, the main difference with respect to the transfer function defined in (3.23) is that a first-order polynomial appears at the numerator of $G(s) = \frac{s}{s + \frac{1}{RC}}$. It is concluded from Figs. 3.1 and 3.7 that the polynomial at the numerator of a transfer function affects time response too, despite the stability not being affected by the numerator of a transfer function.

In this respect, we can use (3.46), the initial value theorem in (3.5), and $V_i(s) = \frac{A}{s}$ to find:

$$\begin{aligned} v_0(0^+) &= \lim_{t \rightarrow 0^+} v_0(t) = \lim_{s \rightarrow \infty} sV_0(s), \\ &= \lim_{s \rightarrow \infty} s \frac{s}{s + \frac{1}{RC}} \frac{A}{s} = A. \end{aligned}$$

The roots of the polynomial at the numerator of a transfer function are called the *transfer function zeros*. Thus, it is concluded that the discontinuity in the time response, when a step input is applied, is due to the fact that the transfer function has

Fig. 3.9 Phase-lead network

the same number of poles and zeros. In the next example, it is shown that a different location of a zero of a transfer function has different effects on the time response.

Example 3.7 Consider the electric circuit shown in Fig. 3.9. This circuit is also known as a *phase-lead network*. The mathematical model of this circuit was obtained in (2.49) and it is rewritten here for ease of reference:

$$V_0(s) = \frac{s+b}{s+a} V_i(s) - \frac{v_c(0)}{s+a}, \quad b = \frac{1}{R_1 C}, \quad a = \frac{1}{C} \frac{R_1 + R_2}{R_1 R_2},$$

where $v_c(0)$ represents the initial voltage at the capacitor and $a > b > 0$. The solution $v_0(t)$ is computed in the following. Substituting $v_i(t) = A$, i.e., $V_i(s) = \frac{A}{s}$:

$$V_0(s) = \frac{(s+b)A}{s(s+a)} - \frac{v_c(0)}{s+a}. \quad (3.47)$$

As $a \neq 0$, the partial fraction expansion results in:

$$\frac{(s+b)A}{s(s+a)} = \frac{B}{s} + \frac{C}{s+a}. \quad (3.48)$$

Constant B is computed as:

$$\begin{aligned} s \frac{(s+b)A}{s(s+a)} \Big|_{s=0} &= s \frac{B}{s} \Big|_{s=0} + s \frac{C}{s+a} \Big|_{s=0}, \\ B &= \frac{bA}{a}. \end{aligned} \quad (3.49)$$

Constant C is computed as:

$$\begin{aligned} (s+a) \frac{(s+b)A}{s(s+a)} \Big|_{s=-a} &= (s+a) \frac{B}{s} \Big|_{s=-a} + (s+a) \frac{C}{s+a} \Big|_{s=-a}, \\ C &= \frac{(b-a)A}{-a}. \end{aligned} \quad (3.50)$$

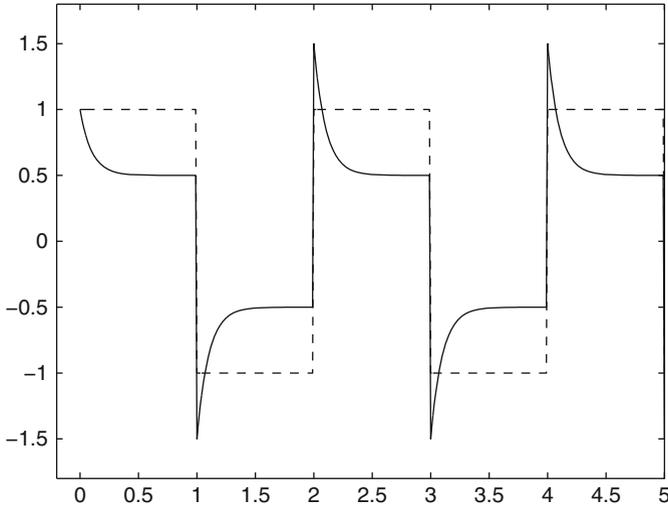


Fig. 3.10 Time response of the circuit in Fig. 3.9. Continuous: $v_0(t)$. Dashed: $v_i(t)$

Replacing (3.48), (3.49), (3.50), in (3.47), it is found that:

$$V_0(s) = \frac{bA}{a} \frac{1}{s} + \frac{(b-a)A}{-a} \frac{1}{s+a} - \frac{v_c(0)}{s+a}.$$

Hence, using the inverse Laplace transform:

$$v_0(t) = \frac{bA}{a} + \left(\frac{(b-a)A}{-a} - v_c(0) \right) e^{-at}. \quad (3.51)$$

This time response is depicted in Fig. 3.10 when $v_i(t)$ is the square wave shown with a dashed line, $b = 5$, $a = 10$ and $v_c(0) = 0$. At $t = 0$, $A = 1$ [V]. Then, (3.51) becomes:

$$v_0(t) = \frac{1}{2} + \frac{1}{2} e^{-10t}.$$

Thus, $v_0(0) = 1$ [V] and $v_0(t_1) \approx \frac{1}{2}$ [V], for $t_1 = 1$ [s], are correctly predicted. Notice that a time interval of 1[s], between $t = 0$ and $t = t_1$, is ten times the time constant of $\frac{1}{a} = 0.1$ [s]; hence, $\frac{1}{2} e^{-10t} \approx 0$ when $t = t_1 = 1$ [s]. On the other hand, applying the KVL to the circuit in Fig. 3.9 it is found that;

$$v_i(t_1) = v_c(t_1) + v_0(t_1),$$

$$v_i(t_1) = A = 1, \quad v_0(t_1) = \frac{R_2}{R_1 + R_2} A,$$

$$v_c(t_1) = v_i(t_1) - v_0(t_1) = \frac{1}{2} = \frac{R_1}{R_1 + R_2}.$$

At $t = t_1 = 1$ [s], $v_i(t)$ changes to $A = -1$. If $t = t - t_1$ is redefined, (3.51) is still valid if it is rewritten as:

$$v_0(t - t_1) = \frac{bA}{a} + \left(\frac{(b-a)A}{-a} - v_c(t_1) \right) e^{-a(t-t_1)},$$

$$v_0(t - t_1) = \frac{-1}{2} + \left(\frac{-1}{2} - \frac{R_1}{R_1 + R_2} \right) e^{-10(t-t_1)},$$

$$v_0(t - t_1) = \frac{-1}{2} - e^{-10(t-t_1)},$$

since $\frac{R_1}{R_1 + R_2} = \frac{1}{2}$. Thus, Fig. 3.10 corroborates that $v_0(t_1) = -1.5$ [V], $t_1 = 1$ [s], and $v_0 \approx -\frac{1}{2}$ [V] at $t = 2$ [s], because $-e^{-10(t-t_1)} \approx 0$ at $t = 2$ [s], are correctly predicted again.

Figure 3.10 has been obtained using the MATLAB/Simulink simulation diagram presented in Fig. 3.11. The *signal generator block* is programmed to have a square wave form, unit amplitude, and 0.5 [Hz] as the frequency. The negative, unit gain block is included because the signal delivered by the signal generator block starts with a negative semicircle. Once the simulation stops, the following MATLAB code is executed in an m-file to draw Fig. 3.10:

```
nn=length(InputOutput(:,1));
n=nn-1;
Ts=5/n;
t=0:Ts:5;
plot(t,InputOutput(:,1),'k--',t,InputOutput(:,2),'k-')
axis([-0.25 5 -2 2])
```

Notice that the transfer function of the circuit in Fig. 3.9, given as:

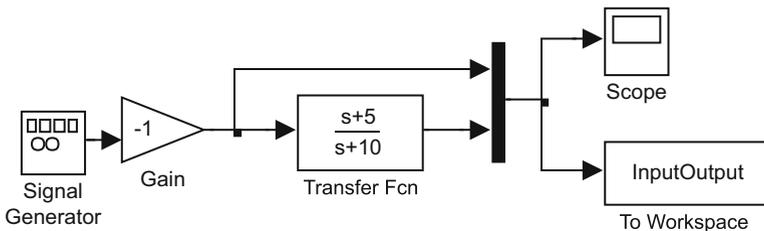


Fig. 3.11 MATLAB/Simulink simulation diagram for circuit in Fig. 3.9

$$\frac{V_0(s)}{V_i(s)} = G(s) = \frac{s+b}{s+a}, \quad (3.52)$$

when $v_c(0) = 0$, has a zero at $s = -b < 0$, whereas the transfer function of the circuit in Fig. 3.6 has a zero at $s = 0$. The main difference produced by this fact is that the time response of the circuit in Fig. 3.6 reaches a zero value in a steady state whereas the steady-state response of the circuit in Fig. 3.9 is given as $\frac{b}{a}A \neq 0$.

Finally, we can use (3.52), the initial value theorem in (3.5), and $V_i(s) = \frac{A}{s}$ to find:

$$\begin{aligned} v_0(0^+) &= \lim_{t \rightarrow 0^+} v_0(t) = \lim_{s \rightarrow \infty} s V_0(s), \\ &= \lim_{s \rightarrow \infty} s \frac{s+b}{s+a} \frac{A}{s} = A. \end{aligned}$$

Thus, it is concluded again that the discontinuity in time response when a step input is applied is due to the fact that the transfer function has the same number of poles and zeros.

3.3 Second-Order Differential Equation

Consider the following linear ordinary second-order differential equation with constant coefficients:

$$\ddot{y} + 2\zeta\omega_n\dot{y} + \omega_n^2y = k\omega_n^2u, \quad y(0) = y_0, \quad \dot{y}(0) = \dot{y}_0, \quad (3.53)$$

where $\omega_n > 0$ and k are real nonzero constants. Assume that $u = A$ is a real constant. Using (3.2) and $U(s) = A/s$, the following can be written:

$$s^2Y(s) - sy_0 - \dot{y}_0 + 2\zeta\omega_n(sY(s) - y_0) + \omega_n^2Y(s) = k\omega_n^2U(s),$$

to obtain:

$$Y(s) = k \frac{\omega_n^2 A}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} + \frac{y_0(s + 2\zeta\omega_n) + \dot{y}_0}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \quad (3.54)$$

First, suppose that all of the initial conditions are zero $\dot{y}_0 = 0$, $y_0 = 0$. Then:

$$Y(s) = k \frac{A \omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s}.$$

Also, suppose that $0 \leq \zeta < 1$ is a real constant, then:

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = (s - a)(s - \bar{a}), \quad (3.55)$$

$$a = \sigma + j\omega_d, \quad \bar{a} = \sigma - j\omega_d, \quad j = \sqrt{-1},$$

$$\omega_d = \omega_n \sqrt{1 - \zeta^2} > 0, \quad \sigma = -\zeta\omega_n \leq 0.$$

Notice that:

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = (s - [\sigma + j\omega_d])(s - [\sigma - j\omega_d]) = (s - \sigma)^2 + \omega_d^2.$$

According to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, in this case the following can be written:

$$Y(s) = k \frac{A\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} = k \frac{A\omega_n^2}{[(s - \sigma)^2 + \omega_d^2]s} = \frac{Bs + C}{(s - \sigma)^2 + \omega_d^2} + \frac{D}{s}, \quad (3.56)$$

where B , C , and D are constants to be computed. The inverse Laplace transform of (3.56) is given as:

$$y(t) = kA \left[1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1 - \zeta^2}} \sin(\omega_d t + \phi) \right], \quad (3.57)$$

$$\phi = \arctan \frac{\sqrt{1 - \zeta^2}}{\zeta}.$$

This expression constitutes the solution of (3.53) when all of the initial conditions are zero, i.e., when $\dot{y}_0 = 0$, $y_0 = 0$. The detailed procedure to find (3.57) from (3.56) is presented in the following. Readers who are not interested in these technical details may continue from (3.62) onward.

Multiplying both sides of (3.56) by the factor $(s - \sigma)^2 + \omega_d^2$ and evaluating at $s = a$ (see (3.55)) yields:

$$k \frac{A\omega_n^2}{[(s - \sigma)^2 + \omega_d^2]s} \Big|_{s=a} = \frac{Bs + C}{(s - \sigma)^2 + \omega_d^2} \Big|_{s=a} + \frac{D}{s} \Big|_{s=a},$$

to obtain:

$$k \frac{A\omega_n^2}{\sigma + j\omega_d} = B(\sigma + j\omega_d) + C.$$

Multiplying the left-hand term by $(\sigma - j\omega_d)/(\sigma - j\omega_d)$:

$$kA\omega_n^2 \frac{(\sigma - j\omega_d)}{\sigma^2 + \omega_d^2} = B\sigma + C + jB\omega_d.$$

Equating the imaginary parts:

$$B = -\frac{kA\omega_n^2}{\sigma^2 + \omega_d^2}. \quad (3.58)$$

Equating the real parts:

$$B\sigma + C = \frac{kA\omega_n^2\sigma}{\sigma^2 + \omega_d^2}.$$

Replacing (3.58):

$$C = \frac{2kA\omega_n^2\sigma}{\sigma^2 + \omega_d^2}. \quad (3.59)$$

Using (3.58) and (3.59), the following can be written:

$$\begin{aligned} \frac{Bs + C}{(s - \sigma)^2 + \omega_d^2} &= \frac{-\frac{kA\omega_n^2}{\sigma^2 + \omega_d^2}s + \frac{2kA\omega_n^2\sigma}{\sigma^2 + \omega_d^2}}{(s - \sigma)^2 + \omega_d^2}, \\ &= -\frac{kA\omega_n^2}{(\sigma^2 + \omega_d^2)} \frac{(s - \sigma)}{[(s - \sigma)^2 + \omega_d^2]} + \frac{kA\omega_n^2\sigma}{(\sigma^2 + \omega_d^2)} \frac{1}{[(s - \sigma)^2 + \omega_d^2]}. \end{aligned}$$

Using the tabulated Laplace transform pairs [4], Ch. 32:

$$\begin{aligned} \mathcal{L}^{-1} \left\{ \frac{s - \sigma}{(s - \sigma)^2 + \omega_d^2} \right\} &= e^{\sigma t} \cos(\omega_d t), \\ \mathcal{L}^{-1} \left\{ \frac{\omega_d}{(s - \sigma)^2 + \omega_d^2} \right\} &= e^{\sigma t} \sin(\omega_d t), \end{aligned}$$

the following is obtained:

$$\mathcal{L}^{-1} \left\{ \frac{Bs + C}{(s - \sigma)^2 + \omega_d^2} \right\} = -\frac{kA\omega_n^2}{\sigma^2 + \omega_d^2} e^{\sigma t} \cos(\omega_d t) + \frac{kA\omega_n^2\sigma}{\omega_d(\sigma^2 + \omega_d^2)} e^{\sigma t} \sin(\omega_d t),$$

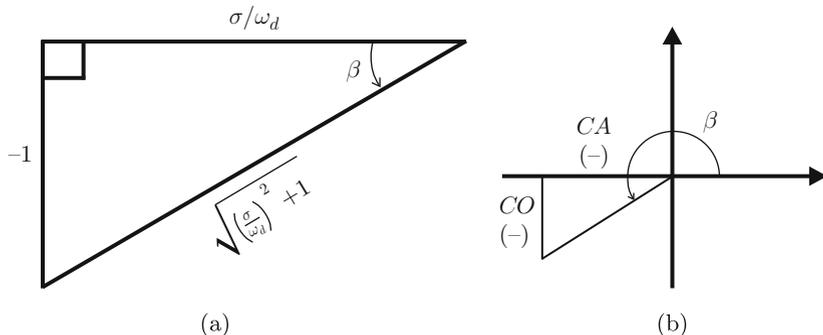


Fig. 3.12 Some important relations for the procedure in Sect. 3.3

$$\begin{aligned}
 &= \frac{kA\omega_n^2}{\sigma^2 + \omega_d^2} e^{\sigma t} \left[-\cos(\omega_d t) + \frac{\sigma}{\omega_d} \sin(\omega_d t) \right], \\
 &= \frac{kA\omega_n^2}{\sigma^2 + \omega_d^2} e^{\sigma t} \left[\sin \beta \cos(\omega_d t) + \cos \beta \sin(\omega_d t) \right] \sqrt{\left(\frac{\sigma}{\omega_d} \right)^2 + 1}.
 \end{aligned}$$

Notice that some relations in Fig. 3.12a have been employed in the last step. On the other hand, it is possible to continue writing:

$$\begin{aligned}
 \sin \beta \cos(\omega_d t) + \cos \beta \sin(\omega_d t) &= \frac{1}{2} [\sin(\beta - \omega_d t) + \sin(\beta + \omega_d t)] + \\
 &\quad + \frac{1}{2} [\sin(\omega_d t - \beta) + \sin(\omega_d t + \beta)], \\
 &= \sin(\omega_d t + \beta),
 \end{aligned}$$

because $\sin(-x) = -\sin(x)$. Hence:

$$\mathcal{L}^{-1} \left\{ \frac{Bs + C}{(s - \sigma)^2 + \omega_d^2} \right\} = kA\omega_n^2 \frac{\sqrt{\left(\frac{\sigma}{\omega_d} \right)^2 + 1}}{\sigma^2 + \omega_d^2} e^{\sigma t} \sin(\omega_d t + \beta).$$

On the other hand, from Fig. 3.12a:

$$\begin{aligned}
 \tan \beta &= \frac{-1}{\sigma/\omega_d} = \frac{-1}{(-\zeta\omega_n)/(\omega_n\sqrt{1-\zeta^2})} = \frac{-1}{\frac{-\zeta}{\sqrt{1-\zeta^2}}}, \\
 \beta &= \pi + \arctan \frac{\sqrt{1-\zeta^2}}{\zeta},
 \end{aligned}$$

where signs of the adjacent and the opposite sides have been taken into account to conclude that β is an angle on the third quadrant (see Fig. 3.12b). This allows the following simplification [4], Ch. 5:

$$\begin{aligned}\sin(\omega_d t + \beta) &= \sin(\omega_d t + \pi + \phi), \\ &= -\sin(\omega_d t + \phi), \\ \phi &= \arctan \frac{\sqrt{1 - \zeta^2}}{\zeta},\end{aligned}$$

and thus:

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{Bs + C}{(s - \sigma)^2 + \omega_d^2}\right\} &= -kA\omega_n^2 \frac{\sqrt{\left(\frac{\sigma}{\omega_d}\right)^2 + 1}}{\sigma^2 + \omega_d^2} e^{\sigma t} \sin(\omega_d t + \phi), \\ &= -kA\omega_n^2 \frac{\sqrt{\sigma^2 + \omega_d^2}}{\omega_d(\sigma^2 + \omega_d^2)} e^{\sigma t} \sin(\omega_d t + \phi), \\ &= -kA\omega_n^2 \frac{1}{\omega_d \sqrt{\sigma^2 + \omega_d^2}} e^{\sigma t} \sin(\omega_d t + \phi), \\ \mathcal{L}^{-1}\left\{\frac{Bs + C}{(s - \sigma)^2 + \omega_d^2}\right\} &= -kA \frac{1}{\sqrt{1 - \zeta^2}} e^{\sigma t} \sin(\omega_d t + \phi),\end{aligned}\quad (3.60)$$

where $\sigma = -\zeta\omega_n$ and $\omega_d = \omega_n\sqrt{1 - \zeta^2}$ have been employed. On the other hand, D in (3.56) is easily computed multiplying both sides of that expression by factor s and evaluating at $s = 0$:

$$D = \frac{kA \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \Big|_{s=0} = kA. \quad (3.61)$$

Using (3.56), (3.60) and (3.61), the following solution is found:

$$y(t) = kA \left[1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1 - \zeta^2}} \sin(\omega_d t + \phi) \right].$$

If it is assumed now that the initial conditions are not zero, then:

$$y(t) = kA \left[1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1 - \zeta^2}} \sin(\omega_d t + \phi) \right] + p(t), \quad (3.62)$$

where:

$$p(t) = \mathcal{L}^{-1} \left\{ \frac{y_0(s + 2\zeta\omega_n) + \dot{y}_0}{s^2 + 2\zeta\omega_n s + \omega_n^2} \right\}. \quad (3.63)$$

The inverse Laplace transform defining $p(t)$ may be computed using a similar procedure to that presented above. Moreover, the reader should realize that $p(t)$ is given by a function as that in (3.60) with coefficients depending on both y_0 and \dot{y}_0 .

The solution given in (3.57) can be decomposed into two parts:

$$y(t) = y_n(t) + y_f(t), \quad (3.64)$$

$$y_n(t) = -kA \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin(\omega_d t + \phi), \quad (3.65)$$

$$y_f(t) = kA, \quad (3.66)$$

where $y_n(t)$ and $y_f(t)$ are the natural and the forced responses respectively. Notice that, according to (3.56) and (3.60), the natural response in (3.65) is due only to the characteristic polynomial $s^2 + 2\zeta\omega_n s + \omega_n^2 = (s - \sigma)^2 + \omega_d^2$, which results from applying the Laplace transform to terms $\ddot{y} + 2\zeta\omega_n \dot{y} + \omega_n^2 y$. Hence, no matter what the input u is, the natural response $y_n(t)$ is given as in (3.65). On the other hand, notice that the forced response in (3.66) is due to term D/s in (3.56), which is introduced by the constant input $u = A$, as $U(s) = A/s$. Also notice that, according to the paragraph after (3.63), when the initial conditions are different from zero they only affect the natural response in the sense that its coefficient depends on the initial conditions y_0 and \dot{y}_0 , but the function of time is always given as $e^{-\zeta\omega_n t} \sin(\omega_d t + \phi)$ or its time derivative. Furthermore, $p(t)$ belongs to the natural response when the initial conditions are different from zero.

The reader may review the procedure presented above after (3.57) to corroborate that, in the case when $-1 < \zeta < 0$ with $k > 0$ and $\omega_n > 0$, the solution in (3.64) or in (3.57) becomes:

$$y(t) = y_n(t) + y_f(t), \quad (3.67)$$

$$y_n(t) = kA \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin \left(\omega_d t - \arctan \left(\frac{\sqrt{1-\zeta^2}}{abs(\zeta)} \right) \right),$$

$$y_f(t) = kA,$$

where $abs(\cdot)$ stands for the absolute value function.

Employing (3.64), when $0 \leq \zeta < 1$, and (3.67), when $-1 < \zeta < 0$, it is concluded that the solution of the differential equation in (3.53) has one of the following behaviors:

1. If $0 < \zeta < 1$, i.e., if both roots of the characteristic polynomial located at $s = -\zeta\omega_n \pm j\omega_d$ have negative real parts $-\zeta\omega_n < 0$, then $\lim_{t \rightarrow \infty} y_n(t) = 0$ and $\lim_{t \rightarrow \infty} y(t) = y_f(t)$.
2. If $-1 < \zeta < 0$, i.e., if both roots of the characteristic polynomial located at $s = -\zeta\omega_n \pm j\omega_d$ have positive real parts $-\zeta\omega_n > 0$, then $y_n(t)$ and $y(t)$ grow without a limit, describing oscillations, as time increases.
3. If $\zeta = 0$, i.e., if both roots of the characteristic polynomial located at $s = -\zeta\omega_n \pm j\omega_d$ have zero real parts $-\zeta\omega_n = 0$, then from (3.64) the following is obtained:

$$\begin{aligned} y(t) &= y_n(t) + y_f(t), \\ y_n(t) &= -kA \cos(\omega_n t), \quad y_f(t) = kA, \end{aligned} \tag{3.68}$$

when the initial conditions are zero, i.e., $y_n(t)$ is an oscillatory function whose amplitude neither increases nor decreases. This means that although $y(t)$ does not grow without a limit, it will not reach the forced response $y_f(t)$ in a steady state.

4. The behavior obtained when ζ is out of the range $-1 < \zeta < 1$ is studied in the subsequent sections as the cases when the roots of the characteristic polynomial are real and repeated or real and different.

Finally, notice that, according to the above discussion and (3.62), (3.63), the solution of a second-order differential equation may present sustained oscillations ($\zeta = 0$) even if the input is zero ($u = 0$) if the initial conditions are different from zero. This explains the oscillations in the electronic circuits studied in Chap. 9 (see last paragraph in Sect. 9.3.1).

3.3.1 Graphical Study of the Solution

The graphical form of solution presented in (3.64) for the differential equation in (3.53) is studied in the following. Recall that it is assumed that $0 \leq \zeta < 1$, $\omega_n > 0$, $k > 0$ and that the initial conditions are zero $y_0 = 0$, $\dot{y}_0 = 0$. Notice that the natural response tends toward zero as time increases if $0 < \zeta < 1$:

$$\lim_{t \rightarrow \infty} y_n(t) = \lim_{t \rightarrow \infty} \left(-kA \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin(\omega_d t + \phi) \right) = 0, \tag{3.69}$$

because $\zeta\omega_n > 0$. Also recall that, in the case when the initial conditions are not zero, the natural response also tends toward zero as time increases. Hence, when time is large, the solution of the differential equation is equal to the forced response:

$$\lim_{t \rightarrow \infty} y(t) = y_f(t) = kA. \tag{3.70}$$

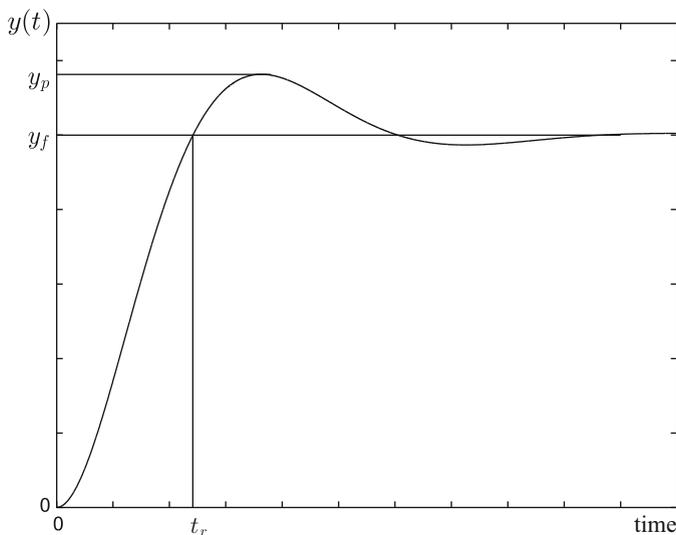


Fig. 3.13 $y(t)$ in (3.64)

This means that the faster the natural response vanishes $y_n(t)$, i.e., as $\zeta\omega_n > 0$ is larger, the faster the complete solution $y(t)$ reaches the forced response $y_f(t)$. As was noted for first-order differential equations, the natural response allows the complete solution $y(t)$ to go from the initial conditions (zero in this case) to the forced response $y_f(t)$.

Two important parameters of the response in (3.64) are the rise time t_r and overshoot M_p , which are shown in Fig. 3.13. Overshoot can be measured as:

$$M_p(\%) = \frac{y_p - y_f}{y_f} \times 100,$$

where $y_f = kA$ stands for the forced response. Recall that it is assumed that $y_0 = \dot{y}_0 = 0$. Carefully analyzing (3.64), it can be shown that [1], pp. 232:

$$t_r = \frac{1}{\omega_d} \left[\pi - \arctan \left(\frac{\sqrt{1 - \zeta^2}}{\zeta} \right) \right], \quad (3.71)$$

$$M_p(\%) = 100 \times e^{-\frac{\zeta}{\sqrt{1 - \zeta^2}} \pi}.$$

Figures 3.14 and 3.15 show how the solution in (3.64) is affected when the parameters ζ and ω_n change. Notice that overshoot M_p is affected only by ζ , whereas rise time t_r is mainly affected by ω_n , although ζ also has a small effect on t_r .

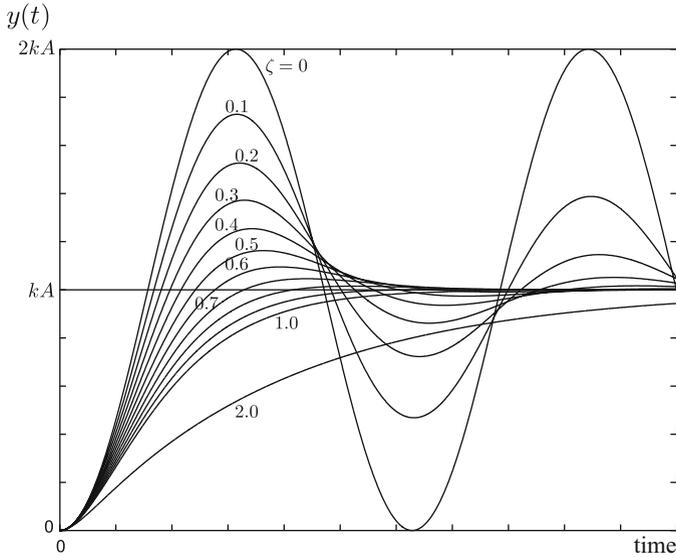


Fig. 3.14 The solution in (3.64) when different values for ζ are used and ω_n is kept constant

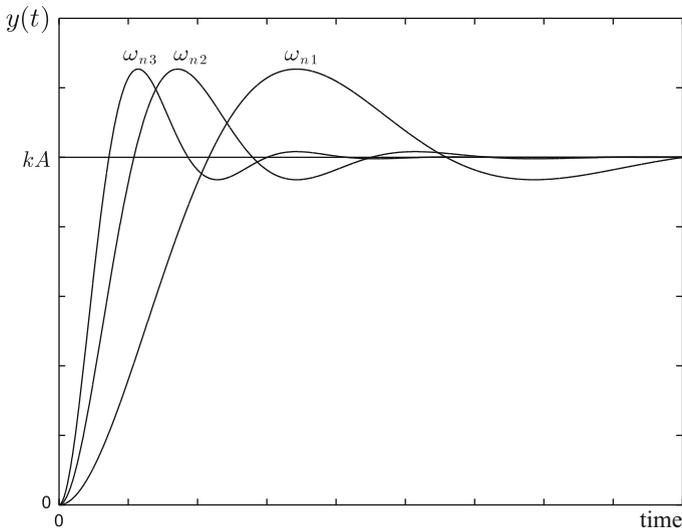


Fig. 3.15 The solution in (3.64) when different values of ω_n are used: $\omega_{n2} = 2\omega_{n1}$, $\omega_{n3} = 3\omega_{n1}$. Parameter ζ is kept constant

Finally, it is important to point out that, in the case when $u = A = 0$ but some of the initial conditions \dot{y}_0 or y_0 are different from zero, $y(t)$ behaves as the time function in (3.65) and is graphically represented as the oscillatory part without the

constant component, i.e., it oscillates around $y = 0$, in any of the Figs. 3.13, 3.14 or 3.15. This is because of function $p(t)$ appearing in (3.62).

3.3.2 Transfer Function

If the initial conditions are zero $y_0 = 0$, $\dot{y}_0 = 0$, then (3.54) can be written as:

$$Y(s) = \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} U(s),$$

or:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad (3.72)$$

where $G(s)$ is the transfer function of the differential equation in (3.53). Notice that $G(s)$, defined in (3.72), has a second-degree characteristic polynomial, i.e., $G(s)$ has two poles. Hence, $G(s)$ is a second-order transfer function. According to (3.55) these poles are located at $s = a$ and $s = \bar{a}$, where:

$$a = \sigma + j\omega_d, \quad \bar{a} = \sigma - j\omega_d.$$

It is important to stress that these poles are complex conjugate only under the condition $-1 < \zeta < 1$. The reader can verify that these poles are real if ζ does not satisfy this condition. Such a case is studied in the subsequent sections.

According to the study in the previous sections, the following can be stated for the output of the transfer function $y(t)$:

- If both poles have a negative real part $\sigma = -\zeta\omega_n < 0$, i.e., $\zeta > 0$, then the natural response $y_n(t)$ vanishes and the complete response $y(t)$ reaches the forced response $y_f(t)$ as time increases. When the input is a constant $u(t) = A$, then the forced response is $y_f(t) = kA$.
- The natural response $y_n(t)$ vanishes faster as $\zeta\omega_n$ is larger. As is depicted in Fig. 3.16, the system response remains enveloped by two exponential functions whose vanishing rate is determined by $\zeta\omega_n$.
- The rise time t_r decreases if ω_n increases. This is corroborated by observing that, according to (3.55), if ω_n increases then ω_d also increases and, according to (3.71), t_r decreases.
- Overshoot M_p decreases if ζ increases.
- According to Fig. 3.17: *i*) the rise time t_r decreases as the complex conjugate poles are located farther to the left of $s = 0$ because ω_n increases (see (3.71) and $\omega_d = \omega_n\sqrt{1 - \zeta^2}$), *ii*) the parameter ζ increases and, hence, overshoot M_p decreases, as the angle θ is larger because $\zeta = \sin(\theta)$, *iii*) $y_n(t)$ vanishes faster as the complex conjugate poles are located farther to the left of $s = 0$ because $\zeta\omega_n$ is larger.

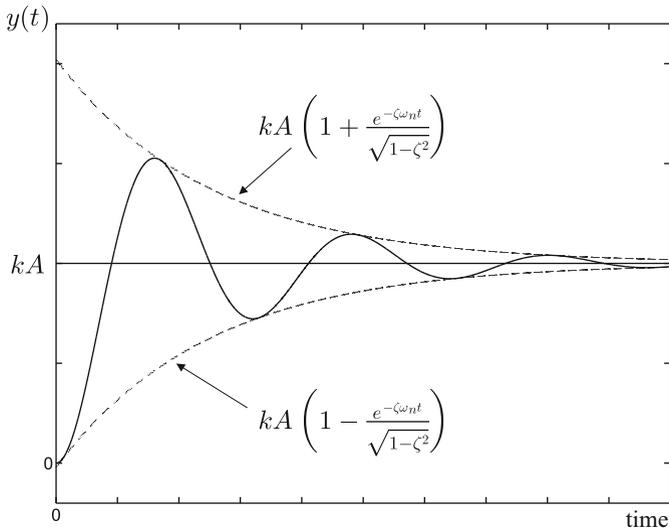
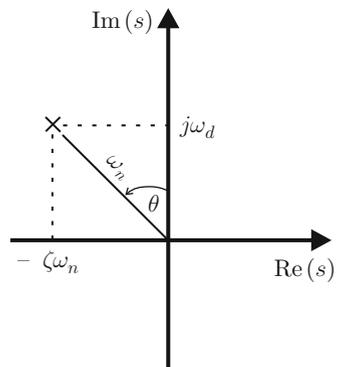


Fig. 3.16 The response of a second-order system is enveloped by two exponential functions when the input is constant and the initial conditions are zero

Fig. 3.17 One pole of $G(s)$ in (3.72) when $0 < \zeta < 1$



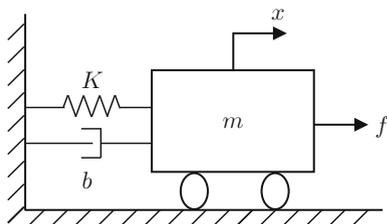
- If $k = 1$ then the transfer function $G(s)$ is said to have unitary gain in a steady state because, when $u = A$ is a constant, the final value theorem (3.4) can be used to find that:

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} sY(s) = \lim_{s \rightarrow 0} s \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \frac{A}{s} = \frac{k\omega_n^2}{\omega_n^2} A = A.$$

The constant k represents the steady-state gain of the transfer function in (3.72).

- If $-1 < \zeta < 0$, then the real part of the poles $\sigma = -\zeta\omega_n$ is positive; hence, the complete response $y(t)$ grows without limit as time increases. Notice that this implies that the complex conjugate poles are located on the right half of the plane s .

Fig. 3.18 A
mass-spring-damper system



- The parameter ζ is known as the damping coefficient because it determines how oscillatory the response $y(t)$ is (see (3.71), for overshoot, and Fig. 3.14).
- The parameter ω_d is known as the damped natural frequency because, according to the previous discussion (see (3.64)), ω_d is the oscillation frequency in $y(t)$ when the damping ζ is different from zero. Notice that the frequency ω_d is the imaginary part of the poles of the transfer function in (3.72) (see (3.55)).
- The parameter ω_n is known as the undamped natural frequency because, according to (3.64) and (3.68), ω_n is the oscillation frequency when damping is zero.

Example 3.8 Consider the mass-spring-damper system depicted in Fig. 3.18. According to Example 2.2, Chap. 2, the mass position x is given by the following second-order differential equation:

$$m\ddot{x} = -Kx - b\dot{x} + f, \quad (3.73)$$

where K is the spring stiffness constant, b is the viscous friction coefficient, m is the body mass, and f is an applied external force. The expression in (3.73) can be written as:

$$\ddot{x} + \frac{b}{m}\dot{x} + \frac{K}{m}x = \frac{1}{m}f,$$

which can be rewritten as the expression in (3.53):

$$\begin{aligned} \ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2x &= k\omega_n^2f, \\ 2\zeta\omega_n &= \frac{b}{m}, \quad \omega_n^2 = \frac{K}{m}, \quad k = \frac{1}{K}, \end{aligned} \quad (3.74)$$

where x is the output y , whereas f is the input u . Hence, $x(t)$ is given as $y(t)$ in (3.64) if $0 \leq \zeta < 1$ when the initial conditions are zero. Notice that all constants in (3.73) are positive: the mass m , the viscous friction coefficient b , and the spring stiffness coefficient K only can be positive. According to (3.74), the damping coefficient is given as:

$$\zeta = \frac{b}{2\sqrt{mK}}. \quad (3.75)$$

If a constant force $f = A$ is applied, the results in Figs. 3.13, 3.14, and 3.15 can be employed to conclude the following.

- If there is no friction, i.e., if $b = 0$, then $\zeta = 0$ and the mass will oscillate forever around the position $x_f = kA = \frac{1}{K}A$. In the case when the external force is zero $f = 0$ but some of the initial conditions \dot{x}_0 or x_0 are different from zero, then the mass will oscillate again, though now around the position $x = 0$, as described by $p(t)$ appearing in (3.62).
- If the friction coefficient $b > 0$ increases, then $\zeta > 0$ also increases and the mass oscillations will disappear because the system will have more damping. It is shown in the subsequent sections that the poles are real and repeated when $\zeta = 1$. In all of these cases, the mass will stop moving at $x = x_f = kA = \frac{1}{K}A$. In the case when the external force is zero $f = 0$ but some initial conditions, \dot{x}_0 or x_0 , are different from zero, the mass will stop at $x = 0$. This is again described by the function $p(t)$ appearing in (3.62). Notice that, according to (3.75), the damping ζ also increases if the mass m or the spring stiffness constant K decreases. To understand the reason for this, consider the opposite case: larger values of both m and K produce larger forces ($m\ddot{x}$ for the mass and Kx for the spring), which, hence, are less affected by the friction force ($b\dot{x}$) and thus, the mass can remain in movement for a longer period of time.
- As the rise time t_r inversely depends on the undamped natural frequency $\omega_n = \sqrt{\frac{K}{m}}$, then a more rigid spring (with a larger K) or a smaller mass produces faster responses (t_r smaller). On the other hand, advantage can be taken of the fact that the oscillation frequency is given by $\omega_d = \omega_n\sqrt{1 - \zeta^2}$ with $\omega_n = \sqrt{\frac{K}{m}}$ to adjust the rate of a mechanical clock based on a mass-spring-damper system: if the clock “lags,” it is necessary to increase its oscillation frequency, or rate, which is achieved by applying tension to the spring to increase K .

The case when $-1 < \zeta < 0$ is not possible in this example because all of the constants b , m , and K are positive. However, the situation when $-1 < \zeta < 0$ commonly occurs in feedback systems as a result of the interconnection of diverse components.

3.4 Arbitrary-Order Differential Equations

In Exercise 3.6, it was shown that some differential equations can be written in terms of time derivatives of both the unknown function $y(t)$ and the excitation function $u(t)$. Moreover, the initial conditions of both variables can be combined to obtain the initial conditions in other circuit components. According to this reasoning, a linear ordinary differential equation with arbitrary order n and constant coefficients can always be written as:

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1\dot{y} + a_0y = b_0u + b_1\dot{u} + \cdots + b_mu^{(m)}, \quad (3.76)$$

where $n \geq m$. If $n < m$ the differential equation has no physical meaning, i.e., there is no physical system in practice whose mathematical model can be represented by such a differential equation and, thus, it is not of any interest for engineers. Notice that all of the derivatives are related to time. Because of this feature, these differential equations are also called *dynamical systems*. The variable y is the unknown variable of the differential equation, whereas u is a function of time that is also called the excitation. The objective of solving a differential equation is to find a function $y(t)$ that satisfies the equality defined by the differential equation. To find $y(t)$, it is necessary to know $u(t)$, the real constants $a_i, i = 0, \dots, n - 1, b_j, j = 0, \dots, m$, and the set of $n + m$ constants $y(0), \dot{y}(0) \dots y^{(n-1)}(0), u(0), \dot{u}(0), \dots, u^{(m-1)}(0)$, which are known as the initial conditions of the problem, representing values that the unknown and the excitation, in addition to their time derivatives, have at $t = 0$.

Applying the Laplace transform (3.2) to (3.76), the following is found:

$$\begin{aligned} s^n Y(s) + a_{n-1} s^{n-1} Y(s) + \dots + a_1 s Y(s) + a_0 Y(s) + P(s) \\ = b_0 U(s) + b_1 s U(s) + \dots + b_m s^m U(s), \end{aligned}$$

where $P(s)$ is a polynomial in s whose coefficients depend on the initial conditions $y(0), \dot{y}(0), \dots, y^{(n-1)}(0), u(0), \dot{u}(0), \dots, u^{(m-1)}(0)$ and the coefficients of the differential equation. Hence, it is possible to write:

$$Y(s) = \frac{b_0 + b_1 s + \dots + b_m s^m}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} U(s) - \frac{P(s)}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}.$$

Suppose that $u(t) = A$ is a constant, i.e., $U(s) = A/s$. Then:

$$\begin{aligned} Y(s) &= \frac{B(s)}{N(s)} \frac{A}{s} - \frac{P(s)}{N(s)}, \\ N(s) &= s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0, \\ B(s) &= b_0 + b_1 s + \dots + b_m s^m. \end{aligned}$$

Also suppose for the moment that all of the initial conditions are zero. This means that $P(s) = 0$ and it is possible to write:

$$Y(s) = \frac{B(s)}{N(s)} \frac{A}{s}.$$

According to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, the solution in time $y(t)$ depends on the roots of the characteristic polynomial $N(s)$, i.e., on poles of the following transfer function:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{B(s)}{N(s)} \quad (3.77)$$

and factor $\frac{1}{s}$ introduced by $U(s) = A/s$. The variable $y(t)$ is known as the *output* and $u(t)$ is known as the *input* of the transfer function $G(s)$. The order of $G(s)$ is defined as the degree of the *characteristic polynomial* $N(s)$. As a n degree polynomial has n roots, then a n order transfer function has n poles (see paragraph after (3.23)). On the other hand, the roots of the polynomial $B(s)$ are known as the *zeros* of the transfer function $G(s)$. If $B(s)$ has the degree m then $G(s)$ has m zeros. Recall the condition $n \geq m$ imposed at the beginning of this section. It is assumed that $B(s)$ and $N(s)$ do not have common roots.

The solution for all of the possible cases for the roots of the polynomial $N(s)$ are studied in the next sections, i.e., all of the possible cases for the poles of $G(s)$. This provides the necessary information to understand how a system of arbitrary order n responds.

3.4.1 Real and Different Roots

Suppose that $N(s)$ has $k \leq n$ real and different roots, which are different from $s = 0$, i.e.,:

$$N(s) = N_0(s)(s - p_1)(s - p_2) \cdots (s - p_k),$$

where $p_i \neq 0$, $i = 1, \dots, k$, with $p_j \neq p_i$ if $i \neq j$, stand for the roots of $N(s)$, whereas $N_0(s)$ is a polynomial containing the remaining $n - k$ roots of $N(s)$, where none of them is located at $s = 0$. According to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, in this case the following must be written:

$$Y(s) = \frac{B(s)A}{N(s)s} = Q(s) + \frac{c_1}{s - p_1} + \frac{c_2}{s - p_2} + \cdots + \frac{c_k}{s - p_k} + \frac{f}{s}, \quad (3.78)$$

where c_i , $i = 1, \dots, k$, and f are real constants to be computed, whereas $Q(s)$ represents all fractions corresponding to the roots of $N_0(s)$. Constant c_i can be computed as follows. Multiply both sides of (3.78) by the factor $(s - p_i)$ and evaluate the resulting expression at $s = p_i$ to obtain:

$$c_i = \left. \frac{B(s)A}{N(s)s} (s - p_i) \right|_{s=p_i}, \quad i = 1, \dots, n.$$

A similar procedure allows us to compute:

$$f = \left. \frac{B(s)A}{N(s)} \right|_{s=0}.$$

Using the tabulated inverse Laplace transform, [4], Ch.32:

$$\mathcal{L}^{-1} \left\{ \frac{c_i}{s - a} \right\} = c_i e^{at},$$

the following is finally obtained:

$$y(t) = q(t) + c_1 e^{p_1 t} + c_2 e^{p_2 t} + \cdots + c_k e^{p_k t} + f,$$

where $q(t) = \mathcal{L}^{-1}\{Q(s)\}$. Notice that the following can be written:

$$\begin{aligned} y(t) &= y_n(t) + y_f(t), \\ y_n(t) &= q(t) + c_1 e^{p_1 t} + c_2 e^{p_2 t} + \cdots + c_k e^{p_k t}, \\ y_f(t) &= f, \end{aligned} \quad (3.79)$$

where the natural response $y_n(t)$ only depends on the roots of the characteristic polynomial $N(s)$, whereas the forced response $y_f(t)$ only depends on the fraction $\frac{1}{s}$ introduced by $U(s)$. Consider the following possibilities.

1. If $p_i < 0$ for all $i = 1, \dots, k$, then $\lim_{t \rightarrow \infty} (c_1 e^{p_1 t} + c_2 e^{p_2 t} + \cdots + c_k e^{p_k t}) = 0$ and $\lim_{t \rightarrow \infty} y(t) = q(t) + y_f(t)$.
2. If $p_i > 0$ for at least one $i = 1, \dots, k$, then $(c_1 e^{p_1 t} + c_2 e^{p_2 t} + \cdots + c_k e^{p_k t}) \rightarrow \infty$, as $t \rightarrow \infty$ and $\lim_{t \rightarrow \infty} y(t) = \infty$.

Example 3.9 Consider the following second-order differential equation also given in (3.53):

$$\ddot{y} + 2\zeta\omega_n\dot{y} + \omega_n^2 y = k\omega_n^2 u, \quad y(0) = y_0, \quad \dot{y}(0) = \dot{y}_0. \quad (3.80)$$

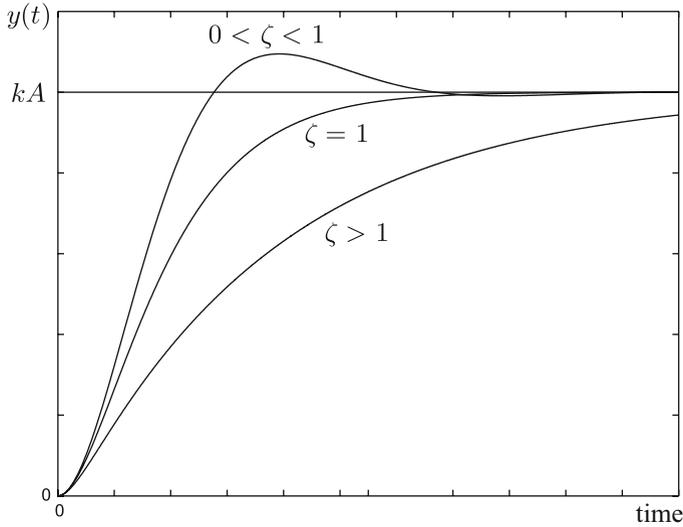
If $\zeta > 1$, then the only two roots, p_1 and p_2 , of the characteristic polynomial $N(s) = s^2 + 2\zeta\omega_n s + \omega_n^2$ are real and different:

$$\begin{aligned} p_1 &= -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}, \\ p_2 &= -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}, \\ p_1 &\neq p_2. \end{aligned}$$

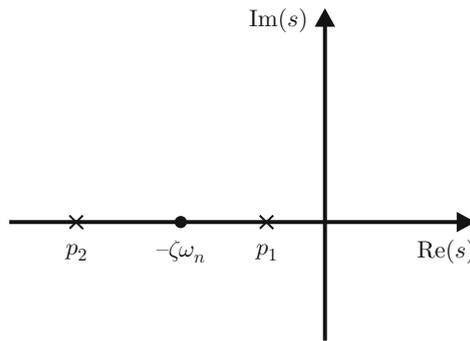
Also, both roots are negative:

$$p_1 < 0, \quad p_2 < 0.$$

Although this is obvious for p_2 , for p_1 , a little observation is required: notice that $-\zeta\omega_n + \omega_n\sqrt{\zeta^2} = 0$, and since $-\zeta\omega_n + \omega_n\sqrt{\zeta^2} > -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}$, then $-\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1} < 0$. In this case $q(t) = 0$, because $Q(s) = 0$ because $N_0(s) = 1$, i.e., $N_0(s)$ has no roots. In Fig. 3.19 the solution $y(t)$ of the differential equation in (3.80) is shown, when $\zeta > 1$, $u = A$ and the initial conditions are zero. Notice that $y(t)$ does not present oscillations in this case. Recall that, when $-1 < \zeta < 1$, the oscillation frequency is equal to the imaginary part of both roots. Hence, when



(a)



(b)

Fig. 3.19 Solution of the second-order differential equation in (3.80), when $\zeta > 1$, and in (3.85), when $\zeta = 1$, with $u = A$ and zero initial conditions . (a) Time response. (b) Location of poles when $\zeta > 1$

$\zeta > 1$, there is no oscillation, because the imaginary part of both roots is zero; thus, the oscillation frequency is also zero.

Example 3.10 Consider again the mass-spring-damper system in Example 3.8. Recall that (3.74) describes the mass movement. Assume that:

$$\zeta = \frac{b}{2\sqrt{mK}} > 1.$$

Then, according to the previous example, the roots p_1 and p_2 of the characteristic polynomial $s^2 + 2\zeta\omega_n s + \omega_n^2$ are real, different, and negative. Notice that this case stands when the viscous friction coefficient $b > 0$ is very large or when the spring stiffness constant $K > 0$ or mass m is very small. Also notice that one root ($p_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1} < 0$) approaches the origin as $\zeta > 1$ becomes larger. According to the discussion above in the present section, this means that the movement of the mass is slower as ζ grows, because it takes more time to function $e^{p_1 t}$ to vanish (despite the function $e^{p_2 t}$ vanishing faster because $p_2 = -\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1} < 0$ moves farther to the left of the origin). This explains why slower responses are depicted in Fig. 3.19, as $\zeta > 1$ is larger.

Example 3.11 Consider the following differential equation:

$$\ddot{y} + c\dot{y} + dy = eu, \quad y(0) = y_0, \quad \dot{y}(0) = \dot{y}_0,$$

with c , d and e some real constants. The roots p_1 and p_2 of the characteristic polynomial $N(s) = s^2 + cs + d$ are given as:

$$p_1 = -\frac{c}{2} + \frac{\sqrt{c^2 - 4d}}{2},$$

$$p_2 = -\frac{c}{2} - \frac{\sqrt{c^2 - 4d}}{2}.$$

The following cases have to be considered:

- If $d < 0$, then:

$$p_1 = -\frac{c}{2} + \frac{\sqrt{c^2 + 4\text{abs}(d)}}{2},$$

$$p_2 = -\frac{c}{2} - \frac{\sqrt{c^2 + 4\text{abs}(d)}}{2},$$

and both roots are real and different, with one positive and the other negative, no matter what the value of c . It is shown, in Example 7.5 of Chap. 7, that this case corresponds to a mass-spring-damper system with a negative spring stiffness constant and that this is obtained in practice when a simple pendulum works around its inverted configuration (also called *unstable*).

- If $c > 0$ and $d > 0$, the case studied in Sect. 3.3 is retrieved when $1 > \zeta > 0$, and the case studied in the present section is retrieved when $\zeta > 1$.
- If $c < 0$ and $d > 0$, with $\text{abs}(c)$ small and d large, the case studied in Sect. 3.3 is retrieved when $-1 < \zeta < 0$. The situation when $\zeta < -1$ is also obtained in this case when $\text{abs}(c)$ is large and d is small, but both roots are positive and different:

$$p_1 = -\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1} > 0,$$

$$p_2 = -\zeta \omega_n - \omega_n \sqrt{\zeta^2 - 1} > 0,$$

$$p_1 \neq p_2.$$

- The cases when $\zeta = -1$ and $\zeta = 1$ are studied in the next section.

3.4.2 Real and Repeated Roots

Suppose that $N(s)$ has one root that is repeated k times and it is different from zero, i.e.,:

$$N(s) = N_0(s)(s - p)^k,$$

with $p \neq 0$ and $N_0(s)$ a polynomial that contains the remaining $n - k$ roots in $N(s)$, and none of them is located at $s = 0$. According to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, the following must be written:

$$Y(s) = \frac{B(s)A}{N(s)s} = Q(s) + \frac{c_k}{(s-p)^k} + \frac{c_{k-1}}{(s-p)^{k-1}} + \cdots + \frac{c_2}{(s-p)^2} + \frac{c_1}{s-p} + \frac{f}{s}, \quad (3.81)$$

where c_i , $i = 1, \dots, k$, and f represent real constants to be computed, whereas $Q(s)$ represents the fractions corresponding to roots of $N_0(s)$. One method of computing c_k is by multiplying both sides of (3.81) by the factor $(s - p)^k$ and evaluating at $s = p$ to obtain:

$$c_k = (s - p)^k \frac{B(s)A}{N(s)s} \Big|_{s=p}.$$

On the other hand, the constants c_i , $i = 1, \dots, k - 1$, can be computed by multiplying both sides of (3.81) by the factor $(s - p)^k$, differentiating $k - i$ times with regard to s and evaluating at $s = p$ to obtain:

$$c_i = \frac{1}{(k - i)!} \frac{d^{k-i}}{ds^{k-i}} \left(\frac{B(s)A}{N(s)s} \right) \Big|_{s=p}, \quad i = 1, \dots, k - 1.$$

Finally, the constant f is computed by multiplying both sides of (3.81) by the factor s and evaluating at $s = 0$ to find:

$$f = \frac{B(s)A}{N(s)} \Big|_{s=0}.$$

Using the tabulated inverse Laplace transform [4], Ch. 32:

$$\mathcal{L}^{-1}\left\{\frac{1}{(s-a)^j}\right\} = \frac{t^{j-1}}{(j-1)!} e^{at}, \quad j = 1, 2, 3, \dots, \quad 0! = 1,$$

the following is found:

$$y(t) = q(t) + c_k \frac{t^{k-1}}{(k-1)!} e^{pt} + c_{k-1} \frac{t^{k-2}}{(k-2)!} e^{pt} + \dots \\ + c_2 t e^{pt} + c_1 e^{pt} + f,$$

where $q(t) = \mathcal{L}^{-1}\{Q(s)\}$. Notice that the following can be written:

$$y(t) = y_n(t) + y_f(t), \\ y_n(t) = q(t) + c_k \frac{t^{k-1}}{(k-1)!} e^{pt} + c_{k-1} \frac{t^{k-2}}{(k-2)!} e^{pt} + \dots \\ + c_2 t e^{pt} + c_1 e^{pt}, \tag{3.82} \\ y_f(t) = f.$$

It is remarked again that $y_n(t)$ only depends on the roots of the characteristic polynomial $N(s)$ and $y_f(t)$ only depends on the fraction $\frac{1}{s}$, introduced by $U(s)$. Consider the following possibilities:

1. If $p < 0$, it is useful to compute the following limit, where j is any positive integer number:

$$\lim_{t \rightarrow \infty} t^j e^{pt} = \lim_{t \rightarrow \infty} \frac{t^j}{e^{-pt}},$$

which represents an indetermination, because $-pt \rightarrow +\infty$. Then, L'Hopital's rule [12], pp. 303, can be employed:

$$\lim_{t \rightarrow \infty} t^j e^{pt} = \lim_{t \rightarrow \infty} \frac{t^j}{e^{-pt}} = \lim_{t \rightarrow \infty} \frac{\frac{dt^j}{dt}}{\frac{de^{-pt}}{dt}} = \lim_{t \rightarrow \infty} \frac{j t^{j-1}}{-p e^{-pt}}.$$

As an indetermination appears again, L'Hopital's rule can be applied several times to obtain:

$$\lim_{t \rightarrow \infty} t^j e^{pt} = \lim_{t \rightarrow \infty} \frac{t^j}{e^{-pt}} = \lim_{t \rightarrow \infty} \frac{\frac{dt^j}{dt}}{\frac{de^{-pt}}{dt}} = \lim_{t \rightarrow \infty} \frac{j t^{j-1}}{-p e^{-pt}},$$

$$= \lim_{t \rightarrow \infty} \frac{j(j-1)t^{j-2}}{(-p)^2 e^{-pt}} = \cdots = \lim_{t \rightarrow \infty} \frac{j!}{(-p)^j e^{-pt}} = 0.$$

Applying this result to the solution obtained above, it is concluded that:

$$\lim_{t \rightarrow \infty} \left(c_k \frac{t^{k-1}}{(k-1)!} e^{pt} + c_{k-1} \frac{t^{k-2}}{(k-2)!} e^{pt} + \cdots + c_2 t e^{pt} + c_1 e^{pt} \right) = 0,$$

and $\lim_{t \rightarrow \infty} y(t) = q(t) + y_f(t)$ for any $p < 0$, no matter how close to zero p is.

2. If $p > 0$, it is clear that:

$$\lim_{t \rightarrow \infty} \left(c_k \frac{t^{k-1}}{(k-1)!} e^{pt} + c_{k-1} \frac{t^{k-2}}{(k-2)!} e^{pt} + \cdots + c_2 t e^{pt} + c_1 e^{pt} \right) \rightarrow \infty,$$

as $t \rightarrow \infty$ and $\lim_{t \rightarrow \infty} y(t) = \infty$.

3. Finally, consider the case when the characteristic polynomial $N(s)$ has k real and repeated roots at $p = 0$, i.e., $N(s) = N_0(s)s^k$, where $N_0(s)$ is a polynomial containing the remaining $n - k$ roots of $N(s)$. Then, according to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, in this case the following must be written:

$$Y(s) = \frac{B(s)A}{N(s)s} = Q(s) + \frac{c_{k+1}}{s^{k+1}} + \frac{c_k}{s^k} + \frac{c_{k-1}}{s^{k-1}} + \cdots + \frac{c_2}{s^2} + \frac{c_1}{s},$$

where $Q(s)$ contains fractions corresponding to roots of $N_0(s)$. Using the tabulated Laplace transform [4], Ch. 32:

$$\mathcal{L}^{-1} \left\{ \frac{1}{s^j} \right\} = \frac{t^{j-1}}{(j-1)!}, \quad j = 1, 2, 3, \dots, \quad 0! = 1, \quad (3.83)$$

yields:

$$y(t) = q(t) + c_{k+1} \frac{t^k}{k!} + c_k \frac{t^{k-1}}{(k-1)!} + \cdots + c_2 t + c_1, \quad (3.84)$$

where $q(t) = \mathcal{L}^{-1}\{Q(s)\}$. In this case:

$$y_n(t) = q(t) + c_k \frac{t^{k-1}}{(k-1)!} + \cdots + c_2 t + c_1,$$

$$y_f(t) = c_{k+1} \frac{t^k}{k!},$$

since the natural response only depends on the k real and repeated roots at $s = 0$ that $N(s)$ has, which, according to (3.83), only introduces terms included in $y_n(t)$. It is clear that $y_n(t) \rightarrow \infty$ and $y(t) \rightarrow \infty$ when $t \rightarrow \infty$ for $k \geq 2$ and that $y_n(t)$ is a constant if $k = 1$. On the other hand, the forced response $y_f(t)$ represents the integral of $u(t)$ iterated k times.

Example 3.12 Consider the following second-order differential equation, also given in (3.53):

$$\ddot{y} + 2\zeta\omega_n\dot{y} + \omega_n^2y = k\omega_n^2u, \quad y(0) = y_0, \quad \dot{y}(0) = \dot{y}_0. \quad (3.85)$$

If $\zeta = 1$, then both roots, p_1 y p_2 , of the characteristic polynomial $N(s) = s^2 + 2\zeta\omega_n s + \omega_n^2$ are real, repeated, and negative:

$$p_1 = p_2 = -\zeta\omega_n.$$

The solution $y(t)$ of the differential equation in (3.85) is depicted in Fig. 3.19, when $\zeta = 1$, $u = A$ and the initial conditions are zero. Notice that $y(t)$ does not exhibit oscillations in this case because both roots have a zero imaginary part. This case, $\zeta = 1$, is at the edge between an oscillatory response $\zeta < 1$ and a response without oscillations $\zeta > 1$. In fact, the case $\zeta = 1$ represents the fastest response without oscillations because, when $\zeta > 1$, the response is slower as ζ increases.

Example 3.13 Consider again the mass-spring-damper system in Example 3.8. Recall that (3.74) describes the movement of mass. Assume that:

$$\zeta = \frac{b}{2\sqrt{mK}} = 1.$$

Then, in this case, the characteristic polynomial $s^2 + 2\zeta\omega_n s + \omega_n^2$ only has a real and negative root:

$$p = -\zeta\omega_n,$$

which is repeated twice. The mass position $x(t)$ evolves exactly as $y(t)$ in Fig. 3.19. This class of response may be very useful if a fast mass movement is required, but without oscillations.

Example 3.14 Suppose that there is neither a spring nor a damper in Fig. 3.18, i.e., assuming that $b = 0$ and $K = 0$, the differential equation in (3.73) becomes:

$$m\ddot{x} = f. \quad (3.86)$$

Notice that in this case the characteristic polynomial is s^2 , which only has a real root at $p = 0$ repeated twice. Now suppose that all of the initial conditions are zero and that the mass receives a small disturbance force described by:

$$f = \begin{cases} \varepsilon_0, & 0 \leq t \leq t_1 \\ 0, & \text{otherwise} \end{cases}, \tag{3.87}$$

where $\varepsilon_0 > 0$ and $t_1 > 0$ are small constant numbers. By direct integration of (3.86), it is found:

$$\dot{x}(t) = \frac{1}{m} \int_0^t f(\tau) d\tau + \dot{x}(0).$$

Again, integrating yields (assuming that $\dot{x}(0) = 0$):

$$x(t) = \frac{1}{m} \int_0^t \left\{ \int_0^r f(\tau) d\tau \right\} dr + x(0),$$

$$x(t) = \frac{1}{m} \int_0^t \left\{ \int_0^r f(\tau) d\tau \right\} dr, \quad x(0) = 0.$$

Replacing (3.87), the following is found:

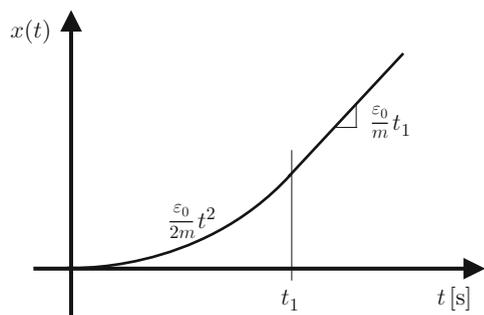
$$x(t) = \begin{cases} \frac{1}{2m} \varepsilon_0 t^2, & 0 \leq t \leq t_1 \\ \frac{1}{2m} \varepsilon_0 t_1^2 + \frac{1}{m} \varepsilon_0 t_1 (t - t_1), & t > t_1 \end{cases}.$$

Notice that $x(t) \rightarrow \infty$ as $t \rightarrow \infty$ (see Fig. 3.20) despite $f = 0$ for $t > t_1$, i.e., despite the forced response being zero for all $t > t_1$. This means that the natural response $x_n(t) \rightarrow \infty$ when $t \rightarrow \infty$, which corroborates results obtained in the present section for real roots at zero, which are repeated at least twice. The reader may resort to everyday experience to verify that a slight hit (f in (3.87)) on the mass m is enough for the mass to begin moving to never stop ($x(t) \rightarrow \infty$) if no friction exists between the mass and the floor.

Example 3.15 Consider the second-order differential equation given in (3.53):

$$\ddot{y} + 2\zeta \omega_n \dot{y} + \omega_n^2 y = k \omega_n^2 u, \quad y(0) = y_0, \quad \dot{y}(0) = \dot{y}_0.$$

Fig. 3.20 Mass position in Fig. 3.18 when it is disturbed and neither a spring nor a damper exists



If $\zeta = -1$, then both roots, p_1 and p_2 , of the characteristic polynomial $N(s) = s^2 + 2\zeta\omega_n s + \omega_n^2$ are real, repeated, and positive:

$$p_1 = p_2 = -\zeta\omega_n > 0.$$

3.4.3 Complex Conjugate and Nonrepeated Roots

Consider the following second-degree polynomial:

$$s^2 + bs + a = (s - p_1)(s - p_2), \quad p_1 = c_1 + jd_1, \quad p_2 = c_2 + jd_2,$$

where p_1 and p_2 are the complex, but not conjugate roots, i.e., c_1, c_2, d_1, d_2 , are real numbers such that $c_1 \neq c_2$ if $d_1 = -d_2$ or $d_1 \neq -d_2$ if $c_1 = c_2$. We have that:

$$\begin{aligned} (s - p_1)(s - p_2) &= s^2 - (p_1 + p_2)s + p_1 p_2, \\ b &= p_1 + p_2 = c_1 + c_2 + j(d_1 + d_2), \\ a &= p_1 p_2 = c_1 c_2 - d_1 d_2 + j(c_2 d_1 + c_1 d_2). \end{aligned}$$

Note that, because of conditions $c_1 \neq c_2$ if $d_1 = -d_2$ or $d_1 \neq -d_2$ if $c_1 = c_2$, one of the coefficients a or b is a complex number, i.e., the polynomial $s^2 + bs + a$ has at least one complex coefficient. In the previous sections, we have seen that physical systems have characteristic polynomials whose coefficients are given only in terms of properties that can be quantified using real numbers, i.e., mass, inertia, friction coefficients, stiffness constant, electrical resistance, inductance, capacitance, etc. Thus, the characteristic polynomials of physical systems cannot have a complex root without including its corresponding complex conjugate. This means that in automatic control we are only interested in characteristic polynomials that have pairs of complex conjugate roots.

Suppose that $N(s)$ has k pairs of complex conjugate and nonrepeated roots, i.e., according to Sect. 3.3, the following can be written:

$$N(s) = N_0(s) \prod_{i=1}^{i=k} [(s - \sigma_i)^2 + \omega_{di}^2],$$

where $\sigma_i \neq \sigma_j$ or $\omega_{di} \neq \omega_{dj}$ if $i \neq j$, whereas $N_0(s)$ is a polynomial containing the remaining $n - 2k$ roots of $N(s)$, none of them located at $s = 0$. Then, according to the partial fraction expansion method [2], Ch. 4, [3], Ch. 7, in this case the following must be written:

$$Y(s) = \frac{B(s)}{N(s)} \frac{A}{s} = Q(s) + \sum_{i=1}^{i=k} \frac{F_i s + C_i}{[(s - \sigma_i)^2 + \omega_{di}^2]} + \frac{f}{s},$$

where $Q(s)$ contains fractions corresponding to roots of $N_0(s)$, f is a constant to be computed as in the previous sections, i.e.,:

$$f = \left. \frac{B(s)A}{N(s)} \right|_{s=0},$$

whereas F_i and C_i , $i = 1, \dots, k$, are constants to be computed as B and C were computed in Sect. 3.3. Hence, using (3.60) the following can be written:

$$y(t) = y_n(t) + y_f(t), \tag{3.88}$$

$$y_n(t) = q(t) + \sum_{i=1}^{i=k} \beta_i e^{-\zeta_i \omega_{ni} t} \sin(\omega_{di} t + \phi_i),$$

$$y_f(t) = f,$$

where β_i , ω_{ni} , ζ_i y ϕ_i , $i = 1, \dots, k$, are real constants (see Sect. 3.3), whereas $q(t) = \mathcal{L}^{-1}\{Q(s)\}$. The evolution of the solution in (3.88) satisfies one of the following cases.

1. If $0 < \zeta_i < 1$ for all $i = 1, \dots, k$, i.e., if the real parts of all of the complex conjugate roots are negative, $-\zeta_i \omega_{ni} < 0$ for all $i = 1, \dots, k$, then $y(t)$ oscillates such that $\lim_{t \rightarrow \infty} \sum_{i=1}^{i=k} \beta_i e^{-\zeta_i \omega_{ni} t} \sin(\omega_{di} t + \phi_i) = 0$ and $\lim_{t \rightarrow \infty} y(t) = q(t) + y_f(t)$.
2. If $-1 < \zeta_i < 0$ for at least one $i = 1, \dots, k$, i.e., if the real part is positive, $-\zeta_i \omega_{ni} > 0$ for at least one of the complex conjugate roots $i = 1, \dots, k$, then $y(t)$ oscillates such that $y_n(t)$ and $y(t)$ grow without a limit as time increases.
3. If $\zeta_i = 0$ for all $i = 1, \dots, k$, i.e., if the real parts of all the complex conjugate and nonrepeated roots are zero, $-\zeta_i \omega_{ni} = 0$ for all $i = 1, \dots, k$, then $\sum_{i=1}^{i=k} \beta_i e^{-\zeta_i \omega_{ni} t} \sin(\omega_{di} t + \phi_i)$ does not disappear as time increases, but nor does it grow without a limit. Thus, although $y(t)$ may not grow without a limit (depending on the behavior of $q(t)$), it does not converge to $y_f(t)$. Notice that for this situation to stand, it is enough that $\zeta_i = 0$ for at least one i and $0 < \zeta_i < 1$ for all of the remaining i .

It is important to point out that, in the cases when $\zeta \geq 1$ or $\zeta \leq -1$, real roots are obtained and one of the cases in Sects. 3.4.1 or 3.4.2 is retrieved.

Example 3.16 The mathematical model of the mechanical system composed of two bodies and three springs depicted in Fig. 3.21 has been obtained in Example 2.4 in Chap. 2. From this result, the case when the spring at the left is not present can be studied. This is achieved by assuming that $K_1 = 0$, i.e.,:

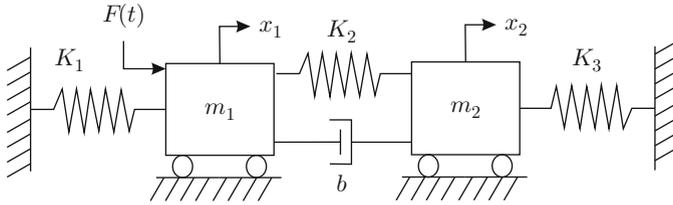


Fig. 3.21 A system with two bodies and three springs

$$\ddot{x}_1 + \frac{b}{m_1}(\dot{x}_1 - \dot{x}_2) + \frac{K_2}{m_1}(x_1 - x_2) = \frac{1}{m_1}F(t),$$

$$\ddot{x}_2 - \frac{b}{m_2}(\dot{x}_1 - \dot{x}_2) + \frac{K_3}{m_2}x_2 - \frac{K_2}{m_2}(x_1 - x_2) = 0.$$

Applying the Laplace transform (3.2) to each one of these differential equations and assuming that all of the initial conditions are zero, the following is found:

$$X_1(s) = \frac{\frac{1}{m_1}F(s) + \left(\frac{b}{m_1}s + \frac{K_2}{m_1}\right)X_2(s)}{s^2 + \frac{b}{m_1}s + \frac{K_2}{m_1}},$$

$$X_2(s) = \frac{\left(\frac{b}{m_2}s + \frac{K_2}{m_2}\right)X_1(s)}{s^2 + \frac{b}{m_2}s + \frac{K_2+K_3}{m_2}}.$$

Replacing the second of these equations in the first one, and rearranging, it is found that:

$$X_1(s) = \frac{\frac{1}{m_1} \left(s^2 + \frac{b}{m_2}s + \frac{K_2+K_3}{m_2} \right) F(s)}{\left(s^2 + \frac{b}{m_1}s + \frac{K_2}{m_1} \right) \left(s^2 + \frac{b}{m_2}s + \frac{K_2+K_3}{m_2} \right) - \left(\frac{b}{m_1}s + \frac{K_2}{m_1} \right) \left(\frac{b}{m_2}s + \frac{K_2}{m_2} \right)}.$$

To simplify the required algebra, assume that $b = 0$, then:

$$X_1(s) = \frac{\frac{1}{m_1} \left(s^2 + \frac{K_2+K_3}{m_2} \right)}{\left(s^2 + \frac{K_2}{m_1} \right) \left(s^2 + \frac{K_2+K_3}{m_2} \right) - \frac{K_2^2}{m_1 m_2}} F(s),$$

or:

$$X_1(s) = \frac{\frac{1}{m_1} \left(s^2 + \frac{K_2+K_3}{m_2} \right)}{s^4 + \left(\frac{K_2}{m_1} + \frac{K_2+K_3}{m_2} \right) s^2 + \frac{K_2 K_3}{m_1 m_2}} F(s). \quad (3.89)$$

Performing the indicated operation, it is found that:

$$s^4 + \left(\frac{K_2}{m_1} + \frac{K_2 + K_3}{m_2} \right) s^2 + \frac{K_2 K_3}{m_1 m_2} = (s^2 + c)(s^2 + e),$$

for some numbers c and e such that:

$$c + e = \frac{K_2}{m_1} + \frac{K_2 + K_3}{m_2}, \quad ce = \frac{K_2 K_3}{m_1 m_2},$$

i.e.,:

$$c = \frac{\frac{K_2}{m_1} + \frac{K_2 + K_3}{m_2} + \sqrt{\left(\frac{K_2}{m_1} - \frac{K_2 + K_3}{m_2} \right)^2 + \frac{4K_2^2}{m_1 m_2}}}{2} > 0,$$

$$e = \frac{\frac{K_2}{m_1} + \frac{K_2 + K_3}{m_2} - \sqrt{\left(\frac{K_2}{m_1} - \frac{K_2 + K_3}{m_2} \right)^2 + \frac{4K_2^2}{m_1 m_2}}}{2} > 0.$$

Then, the following can be written:

$$X_1(s) = \frac{1}{m_1} \left(s^2 + \frac{K_2 + K_3}{m_2} \right) F(s). \quad (3.90)$$

As e and c are positive and different, it is ensured that the characteristic polynomial in (3.90) has two pairs of complex conjugate and different roots (in fact, the roots are imaginary conjugate and different). Hence, mass m_1 , whose position is represented by $x_1(t)$, moves by oscillating and it never stops. However, the amplitude of this oscillatory movement does not grow without a limit. These oscillations are performed at two angular frequencies defined by $\omega_{d1} = \omega_{n1} = \sqrt{c}$ and $\omega_{d2} = \omega_{n2} = \sqrt{e}$. Recall that the damping of these oscillations is zero.

3.4.4 Complex Conjugated and Repeated Roots

Assume that the characteristic polynomial $N(s)$ has a pair of complex conjugate roots that are repeated $k > 1$ times, i.e.,:

$$N(s) = N_0(s)(s^2 + 2\zeta\omega_n s + \omega_n^2)^k = N_0(s)[(s - \sigma)^2 + \omega_d^2]^k,$$

where $N_0(s)$ is a polynomial that includes the remaining $n - 2k$ roots of $N(s)$, without any of them located at $s = 0$. Then, the partial fraction expansion method establishes that in this case [2], Ch. 4:

$$Y(s) = \frac{B(s)}{N(s)} \frac{A}{s} = Q(s) + \frac{F_1 s + C_1}{[(s - \sigma)^2 + \omega_d^2]^k} + \frac{F_2 s + C_2}{[(s - \sigma)^2 + \omega_d^2]^{k-1}} + \cdots \\ + \frac{F_{k-1} s + C_{k-1}}{[(s - \sigma)^2 + \omega_d^2]^2} + \frac{F_k s + C_k}{(s - \sigma)^2 + \omega_d^2} + \frac{f}{s},$$

where $Q(s)$ contains fractions corresponding to the roots of $N_0(s)$. In the case of real roots, it was found that when roots are not repeated, they introduce functions such as e^{pt} , where p is the corresponding root. When such a root is repeated j times, it was found that the functions introduced become $t^{j-1} e^{pt}$. In the case of complex conjugate and repeated roots, a similar situation occurs. In this case, any formal proof is not presented because of cumbersome notation, and only intuitive ideas are given to understand the reason for the resulting functions.

Using the inverse Laplace transform in (3.60), it was found that a complex conjugate nonrepeated root introduces the time function:

$$\mathcal{L}^{-1} \left\{ \frac{Bs + C}{(s - \sigma)^2 + \omega_d^2} \right\} = \beta e^{\sigma t} \sin(\omega_d t + \phi),$$

for some constants β and ϕ . Then, in the case of complex conjugated and repeated roots, $y(t)$ is given as:

$$y(t) = y_n(t) + y_f(t), \tag{3.91} \\ y_n(t) = q(t) + \beta_k t^{k-1} e^{\sigma t} \sin(\omega_d t + \phi_k) + \beta_{k-1} t^{k-2} e^{\sigma t} \sin(\omega_d t + \phi_{k-1}) \\ + \cdots + \beta_2 t e^{\sigma t} \sin(\omega_d t + \phi_2) + \beta_1 e^{\sigma t} \sin(\omega_d t + \phi_1), \\ y_f(t) = f,$$

where $q(t) = \mathcal{L}^{-1}\{Q(s)\}$. The solution in (3.91) has one of the following behaviors.

1. If the complex conjugate root has a positive real part $\sigma = -\zeta \omega_n > 0$, i.e., if $-1 < \zeta < 0$, it is easy to see that $y_n(t) \rightarrow \infty$ and $y(t) \rightarrow \infty$ as $t \rightarrow \infty$.
2. If the complex conjugate root has a negative real part $\sigma = -\zeta \omega_n < 0$, i.e., if $0 < \zeta < 1$, it is possible to proceed as in Sect. 3.4.2 to compute the following limit:

$$\lim_{t \rightarrow \infty} t^j e^{\sigma t} \sin(\omega_d t + \phi) = 0,$$

for any positive integer j and any real negative σ . Then, for complex conjugate and repeated roots with a negative real part:

$$\lim_{t \rightarrow \infty} \left(\beta_k t^{k-1} e^{\sigma t} \sin(\omega_d t + \phi_k) + \beta_{k-1} t^{k-2} e^{\sigma t} \sin(\omega_d t + \phi_{k-1}) \right. \\ \left. + \cdots + \beta_2 t e^{\sigma t} \sin(\omega_d t + \phi_2) + \beta_1 e^{\sigma t} \sin(\omega_d t + \phi_1) \right) = 0, \tag{3.92}$$

and $\lim_{t \rightarrow \infty} y(t) = q(t) + y_f(t)$.

3. If the root has a zero real part, i.e., $\sigma = 0$, then:

$$\begin{aligned} y(t) &= y_n(t) + y_f(t), \\ y_n(t) &= q(t) + \beta_k t^{k-1} \sin(\omega_d t + \phi_k) + \beta_{k-1} t^{k-2} \sin(\omega_d t + \phi_{k-1}) + \cdots \\ &\quad + \beta_2 t \sin(\omega_d t + \phi_2) + \beta_1 \sin(\omega_d t + \phi_1). \\ y_f(t) &= f. \end{aligned}$$

Notice that, in this case:

$$\begin{aligned} &\beta_k t^{k-1} \sin(\omega_d t + \phi_k) + \beta_{k-1} t^{k-2} \sin(\omega_d t + \phi_{k-1}) + \cdots \\ &+ \beta_2 t \sin(\omega_d t + \phi_2) + \beta_1 \sin(\omega_d t + \phi_1) \rightarrow \infty, \end{aligned} \quad (3.93)$$

as time increases in the case when the root is repeated at least twice, i.e., if $k > 1$. But if this root is not repeated, i.e., if $k = 1$, then:

$$\begin{aligned} &\beta_k t^{k-1} \sin(\omega_d t + \phi_k) + \beta_{k-1} t^{k-2} \sin(\omega_d t + \phi_{k-1}) + \cdots \\ &+ \beta_2 t \sin(\omega_d t + \phi_2) + \beta_1 \sin(\omega_d t + \phi_1) = \beta_1 \sin(\omega_d t + \phi_1), \end{aligned} \quad (3.94)$$

is an oscillatory function whose amplitude neither grows nor vanishes, i.e., although $y(t)$ does not grow without a limit (depending on $q(t)$) it will never reach $y_f(t)$ in a steady state.

3.4.5 Conclusions

In the previous sections it has been assumed that all of the initial conditions are zero, i.e., that:

$$-\frac{P(s)}{N(s)} = 0.$$

If some of the initial conditions are not zero, then the previous expression is not valid and the solutions $y(t)$ found in the previous sections have to be completed adding the function:

$$w(t) = \mathcal{L}^{-1} \left\{ -\frac{P(s)}{N(s)} \right\}.$$

As the denominator of:

$$W(s) = \mathcal{L}\{w(t)\} = -\frac{P(s)}{N(s)},$$

is still the characteristic polynomial $N(s)$, then, according to the partial fraction expansion method, $w(t)$ contains the same functions of time contained in the natural response $y_n(t)$ found in every case studied in the previous sections. This means that the complete natural response in every one of the cases studied in the previous sections must also include $w(t)$.

It is said that the differential equation in (3.76) or, equivalently, that transfer function $G(s) = \frac{B(s)}{N(s)}$, is:

- *Stable* if the natural response tends toward zero as time increases.
- *Unstable* if the natural response grows without a limit as time increases.
- *Marginally stable* if the natural response neither grows without limit nor tends toward zero as time increases.

Assume that all of the cases studied in the previous sections, i.e., regarding the roots of $N(s)$, appear simultaneously. Recall that the roots of $N(s)$ represent the poles of $G(s)$. The following can be concluded:

Conditions for the Stability of a Transfer Function

1. If all poles of $G(s)$ have negative real parts then $G(s)$ is stable.
2. If all poles of $G(s)$ have negative real parts, except for some nonrepeated poles having zero real parts then the transfer function $G(s)$ is marginally stable.
3. If at least one pole of $G(s)$ has a positive real part, then the transfer function $G(s)$ is unstable.
4. If there is at least one pole of $G(s)$ with a zero real part which is repeated at least twice, then $G(s)$ is unstable.

On the other hand, it has also been seen that $y_f(t)$ depends on the input $u(t)$ and that, in fact, both of them are represented by “similar” time functions if the characteristic polynomial has no root at $s = 0$, i.e., if $G(s)$ has no pole at $s = 0$. This means that in a control system, the input variable $u(t)$ can be used as the value it is desired that the output $y(t)$ reaches, i.e., $u(t)$ can be used to specify the desired value for $y(t)$. This is accomplished as follows. If the transfer function is stable, i.e., if $y_n(t) \rightarrow 0$, then $y(t) \rightarrow y_f(t)$; hence, the only thing that remains is to ensure that $y_f(t) = u$. The conditions to achieve this when $u = A$ is a constant are established in the following.

Conditions to Ensure $\lim_{t \rightarrow \infty} y(t) = A$

The complete response $y(t)$ reaches the desired output $u = A$ as time increases, i.e., $y_f(t) = A$ and $y_n(t) \rightarrow 0$, if $G(s)$ is stable and the coefficients of the independent terms of the polynomials $B(s)$ and $N(s)$ are equal, i.e., if $B(0) = N(0)$. This can be verified using the final value theorem (3.4):

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} sY(s) = \lim_{s \rightarrow 0} s \frac{B(s)}{N(s)} \frac{A}{s} = \frac{B(0)}{N(0)} A = A. \quad (3.95)$$

Under these conditions, it is said that $G(s)$ is a transfer function with unitary gain in a steady state. When u is not a constant and it is desired that $\lim_{t \rightarrow \infty} y(t) = u(t)$,

$G(s)$ is also required to be stable, i.e., that $y_n(t) \rightarrow 0$, but some additional conditions must be satisfied. The determination of these conditions and how to satisfy them is the subject of the study presented in the subsequent chapters of this book (see Sect. 4.4).

Example 3.17 Consider the situation presented in Example 3.14. From this example, an experiment can be designed that allows us to know whether a system, a differential equation or a transfer function is stable or unstable:

If the system is originally at “rest,” apply a pulse disturbance to it and observe its evolution:

- *If the system is stable then it “moves” and after a while it comes back to the configuration where it was originally at “rest.”*
- *If the system is unstable, then it “moves” more and more such that it goes far away from the configuration where it was originally at “rest.”*

Example 3.18 Consider the mass-spring-damper system studied in Example 3.8, but now assume that the spring is not present. Replacing $K = 0$ in (3.73) yields:

$$m\ddot{x} + b\dot{x} = f.$$

Suppose that a force is applied according to:

$$f = k_p(x_d - x), \quad (3.96)$$

where k_p is a positive constant, x_d is a constant standing for the desired position, and x is the mass position. Combining these expressions yields:

$$m\ddot{x} + b\dot{x} = k_p(x_d - x),$$

and rearranging:

$$\ddot{x} + \frac{b}{m}\dot{x} + \frac{k_p}{m}x = \frac{k_p}{m}x_d.$$

Using the Laplace transform and assuming that all of the initial conditions are zero, the following is found:

$$G(s) = \frac{X(s)}{X_d(s)} = \frac{\frac{k_p}{m}}{s^2 + \frac{b}{m}s + \frac{k_p}{m}}.$$

As an exercise, the reader is to find the roots of the characteristic polynomial $s^2 + \frac{b}{m}s + \frac{k_p}{m}$ to corroborate that both roots have a negative real part if $\frac{b}{m} > 0$ and $\frac{k_p}{m} > 0$. Then, using (3.95), it is found that the position reached in a steady state:

$$\lim_{t \rightarrow \infty} x(t) = \lim_{s \rightarrow 0} sX(s) = \lim_{s \rightarrow 0} s \frac{\frac{k_p}{m}}{s^2 + \frac{b}{m}s + \frac{k_p}{m}} \frac{x_d}{s} = \frac{k_p}{m} x_d = x_d,$$

equals the desired position x_d . The reason for this result can be explained using, again, everyday experience: when $x = x_d$, the force produced according to (3.96) is $f = 0$; hence, the mass can stop and remain at that position. Moreover, if $x < x_d$, then $f > 0$ and the mass increases its position x approaching x_d (see Fig. 3.18 to recall that x increases in the same direction where f is positive). If $x > x_d$, then $f < 0$ and the mass position x decreases, approaching x_d again.

As an exercise, the reader is to verify that, in the case where a spring is present, i.e., when $K > 0$, then $\lim_{t \rightarrow \infty} x(t) \neq x_d$ if $x_d \neq 0$. Notice, however, that again this can be explained using everyday experience:

If $x = x_d$ then, according to (3.96), the external force applied to the mass is zero $f = 0$ and the friction force is also zero $-b\dot{x} = 0$, if it is assumed that the mass is at rest. However, if $x = x_d \neq 0$ then the force of the spring on the mass $-Kx$ is different from zero; hence, the mass will abandon the position $x = x_d \neq 0$. The mass will reach a position x such that the force of spring and the force f in (3.96) exactly cancel each other out, i.e., where $k_p(x_d - x) = Kx$.

3.5 Poles and Zeros in Higher-Order Systems

Systems of orders greater than two are known as higher-order systems. Contrary to first and second order systems, in higher-order systems it is not possible to perform a detailed graphical study of the solution $y(t)$. The main reason for this is the complexity of the expressions arising when the transfer function has three or more poles. Hence, it is important to approximate a higher-order system using a system with a smaller order. There are two ways of accomplishing this: *i*) The approximate cancellation of some poles with some zeros of the corresponding transfer function, and *ii*) Neglecting the effect of “fast” poles. Some examples are presented in the following.

3.5.1 Approximate Pole-Zero Cancellation and Reduced Order Models

Consider the following second-order system:

$$Y(s) = \frac{k(s-d)}{(s-p_1)(s-p_2)} U(s). \quad (3.97)$$

Suppose that $U(s) = A/s$, $p_1 \neq p_2$, $p_1 < 0$, $p_2 < 0$, $d < 0$, and $p_1 p_2 \approx -kd$ to render approximately the unitary gain in a steady state of the transfer function in (3.97). Using partial fraction expansion:

$$Y(s) = \frac{k(s-d)}{(s-p_1)(s-p_2)} \frac{A}{s} = \frac{B}{s-p_1} + \frac{C}{s-p_2} + \frac{D}{s}. \quad (3.98)$$

Multiplying both sides by the factor $(s-p_1)$ and evaluating at $s = p_1$ yields:

$$B = \left. \frac{k(s-d)A}{(s-p_2)s} \right|_{s=p_1} = \frac{k(p_1-d)A}{(p_1-p_2)p_1}.$$

Multiplying both sides of (3.98) by the factor $(s-p_2)$ and evaluating at $s = p_2$ yields:

$$C = \left. \frac{k(s-d)A}{(s-p_1)s} \right|_{s=p_2} = \frac{k(p_2-d)A}{(p_2-p_1)p_2}.$$

Multiplying both sides of (3.98) by the factor s and evaluating at $s = 0$ yields:

$$D = \left. \frac{k(s-d)A}{(s-p_1)(s-p_2)} \right|_{s=0} = -\frac{kdA}{p_1 p_2}.$$

If $p_1 \approx d < 0$ then, according to $p_1 p_2 \approx -kd$, we have $k \approx -p_2$ and, hence:

$$\begin{aligned} B &\approx 0, \\ C &\approx \frac{k(p_2-d)A}{(p_2-d)p_2} = \frac{kA}{p_2}, \\ D &\approx -\frac{kA}{p_2}, \end{aligned}$$

to finally obtain:

$$y(t) \approx \frac{kA}{p_2} e^{p_2 t} - \frac{kA}{p_2}.$$

The reader can verify that:

$$y(t) = \frac{kA}{p_2} e^{p_2 t} - \frac{kA}{p_2},$$

is the solution of:

$$Y(s) = \frac{k}{s-p_2} U(s), \quad (3.99)$$

with $U(s) = A/s$ and $k = -p_2$. Thus, the following is concluded. If one pole and one zero of a transfer function are very close, then they approximately cancel each other out to obtain a reduced order transfer function. This means that (3.99) can be used instead of (3.97) to obtain very similar results. The advantages of using the model in (3.99) are: *i*) it is a model with a smaller order than (3.97), and *ii*) it has no zero. The advantage of using a reduced order model, which approximately describes a higher-order system is that the response of a reduced order system can often be handled as that of a first- or a second-order system. On the other hand, a zero in a transfer function modifies the time response in a manner that is not easy to quantify; hence, a transfer function that has no zeros is preferable.

It is important to say that the simplification obtained by the cancellation of one pole and one zero is useful only if the pole and the zero have negative real parts. This is because one pole with a positive real part, which is not exactly cancelled because of parameter uncertainty, has a dangerous effect that becomes larger as time increases.

3.5.2 Dominant Poles and Reduced Order Models

Consider the following transfer function:

$$Y(s) = \frac{k}{s + \frac{1}{RC}} \frac{d}{s + a} U(s), \quad (3.100)$$

where $U(s) = A/s$. Using partial fraction expansion, the following is found:

$$Y(s) = \frac{k}{s + \frac{1}{RC}} \frac{d}{s + a} \frac{A}{s} = \frac{B}{s + \frac{1}{RC}} + \frac{C}{s + a} + \frac{D}{s}. \quad (3.101)$$

Multiplying both sides by the factor $s + \frac{1}{RC}$ and evaluating at $s = -\frac{1}{RC}$:

$$B = k \frac{d}{s + a} \frac{A}{s} \Big|_{s = -\frac{1}{RC}} = k \frac{d}{a - \frac{1}{RC}} \frac{A}{\left(-\frac{1}{RC}\right)}.$$

Multiplying both sides of (3.101) by the factor $s + a$ and evaluating at $s = -a$:

$$C = k \frac{d}{s + \frac{1}{RC}} \frac{A}{s} \Big|_{s = -a} = k \frac{d}{-a + \frac{1}{RC}} \frac{A}{(-a)}.$$

Multiplying both sides of (3.101) by the factor s and evaluating at $s = 0$:

$$D = k \frac{d}{s + \frac{1}{RC}} \frac{A}{(s + a)} \Big|_{s = 0} = k \frac{d}{\frac{1}{RC}} \frac{A}{a}.$$

If $a \gg \frac{1}{RC} > 0$, then:

$$B \approx -k \frac{d}{a} \frac{A}{\frac{1}{RC}},$$

$$C \approx k \frac{dA}{a^2} \ll D.$$

Hence, using (3.101) and neglecting C , the following is found:

$$y(t) \approx -k \frac{d}{a} \frac{A}{\frac{1}{RC}} e^{-\frac{1}{RC}t} + k \frac{d}{\frac{1}{RC}} \frac{A}{a}.$$

The reader can verify that:

$$y(t) = -k \frac{d}{a} \frac{A}{\frac{1}{RC}} e^{-\frac{1}{RC}t} + k \frac{d}{\frac{1}{RC}} \frac{A}{a},$$

is the solution of:

$$Y(s) = \frac{kd}{a(s + \frac{1}{RC})} U(s), \quad (3.102)$$

with $U(s) = A/s$. Thus, (3.102) can be used instead of (3.100). The condition $a \gg \frac{1}{RC}$ is interpreted as “the pole at $s = -a$ is very fast compared with the pole at $s = -\frac{1}{RC}$ ”. The pole at $s = -\frac{1}{RC}$ is known as the dominant pole because its effect is more important in the system response. The following is concluded. If one pole is much faster than the others, then a reduced order transfer function can be obtained by neglecting the fastest pole to only maintain the dominant (the slowest) poles. An accepted criterion is that the fastest poles (those to be neglected) have a real part that is five times the real parts of the dominant poles (those to be kept). Notice, however, that it is not just a matter of simply eliminating the factor $s + a$ at the denominator of the transfer function: the constant a still appears at the denominator of (3.102) because a is necessary to keep without change the steady-state gain of the transfer function in (3.100).

A simple way of obtaining (3.102) from (3.100) is the following:

$$Y(s) = \frac{k}{s + \frac{1}{RC}} \frac{d}{s + a} U(s) = \frac{k}{s + \frac{1}{RC}} \frac{d}{a(\frac{1}{a}s + 1)} U(s). \quad (3.103)$$

If a is very large it can be assumed that $\frac{1}{a}s \ll 1$; hence:

$$Y(s) \approx \frac{kd}{a(s + \frac{1}{RC})} U(s),$$

which is the expression in (3.102).

Finally, it is important to say that order reduction, as presented above, is only valid if the neglected pole has a negative real part. If the neglected pole has a positive real part, then the contribution of this pole will grow to infinity as time increases, no matter how small C is. Thus, the contribution of such a pole cannot be neglected.

Example 3.19 According to Exercise 9 and Example 2.20 in Chap. 2, the mathematical model of a permanent magnet brushed DC motor is given as:

$$\begin{aligned} L_a \frac{di_a}{dt} &= v - R_a i_a - k_e \omega, \\ J \frac{d\omega}{dt} &= k_m i_a - B\omega, \end{aligned} \quad (3.104)$$

where ω is the motor velocity (also see (10.9) and (10.10) in Chap. 10). Suppose that this motor actuates on a hydraulic mechanism, which, by centrifugal effect, produces a water flow, q_i , which is proportional to the motor velocity, i.e., $q_i = \gamma \omega$, where γ is a positive constant. Finally, this water flow enters the tank in Example 3.2 in the present chapter, whose mathematical model is given in (3.28), rewritten here for ease of reference:

$$\begin{aligned} \frac{dh}{dt} + ah &= kq_i \\ a &= \frac{1}{RC}, \quad k = \frac{1}{C}. \end{aligned}$$

Using the Laplace transform (3.2) and assuming that all the initial conditions are zero, it is not difficult to verify that the corresponding equations become:

$$I(s) = \frac{1/L_a}{s + \frac{R_a}{L_a}} (V(s) - k_e \omega(s)), \quad (3.105)$$

$$\omega(s) = \frac{k_m/J}{s + \frac{B}{J}} I(s), \quad (3.106)$$

$$H(s) = \frac{1/C}{s + \frac{1}{RC}} Q_i(s). \quad (3.107)$$

Combining (3.105) and (3.106) yields:

$$\omega(s) = \frac{k_m/J}{s + \frac{B}{J}} \frac{1/L_a}{s + \frac{R_a}{L_a}} (V(s) - k_e \omega(s)).$$

According to (3.103), the following can be written:

$$\omega(s) = \frac{k_m/J}{\frac{B}{J} \left(\frac{J}{B}s + 1 \right)} \frac{1/L_a}{\frac{R_a}{L_a} \left(\frac{L_a}{R_a}s + 1 \right)} (V(s) - k_e \omega(s)).$$

In small motors, the inductance L_a and the viscous friction coefficient B are small; hence, $\frac{L_a}{R_a} \ll \frac{J}{B}$. Thus, it can be assumed that $\frac{L_a}{R_a}s \ll 1$ and the following can be written:

$$\omega(s) = \frac{1}{s + \frac{B}{J}} \frac{k_m}{JR_a} (V(s) - k_e \omega(s)).$$

Rearranging terms, this expression can be rewritten as:

$$\omega(s) = \frac{\frac{k_m}{JR}}{s + \left(\frac{B}{J} + \frac{k_m k_e}{JR_a}\right)} V(s).$$

Replacing this and $Q_i(s) = \gamma \omega(s)$ in (3.107) yields:

$$H(s) = \frac{\frac{1}{C}}{s + \frac{1}{RC}} \gamma \frac{\frac{k_m}{JR}}{s + \left(\frac{B}{J} + \frac{k_m k_e}{JR_a}\right)} V(s).$$

Proceeding again as in (3.103), the following can be written:

$$H(s) = \frac{\frac{\gamma}{C}}{\frac{1}{RC} (RCs + 1)} \frac{\frac{k_m}{JR}}{\left(\frac{B}{J} + \frac{k_m k_e}{JR_a}\right) \left(\frac{1}{\frac{B}{J} + \frac{k_m k_e}{JR_a}} s + 1\right)} V(s).$$

If the tank cross-section is large enough, the motor will reach its nominal velocity a long time before the water level in tank increases significantly. This can be stated by saying that “the tank time constant is very large compared with the motor time constant,” i.e., $RC \gg \frac{1}{\frac{B}{J} + \frac{k_m k_e}{JR_a}}$. Then, it is possible to say that $\frac{1}{\frac{B}{J} + \frac{k_m k_e}{JR_a}} s \ll 1$; hence, the following approximation is valid:

$$H(s) = \frac{\frac{\gamma}{C}}{\frac{1}{RC} (RCs + 1)} \frac{\frac{k_m}{JR}}{\left(\frac{B}{J} + \frac{k_m k_e}{JR_a}\right)} V(s),$$

or:

$$H(s) = \frac{\frac{1}{C}}{s + \frac{1}{RC}} \frac{\frac{k_m}{JR} \gamma}{\left(\frac{B}{J} + \frac{k_m k_e}{JR_a}\right)} V(s).$$

Then, defining:

$$k_1 = \frac{\frac{k_m}{JR} \gamma}{\frac{B}{J} + \frac{k_m k_e}{JR_a}},$$

and comparing with (3.107), the expression in (3.139) is justified, i.e., that the water flow is proportional to the voltage applied to the motor through a constant k_1 . This is possible if: 1) the motor electrical time constant is smaller than the motor mechanical time constant, i.e., if $\frac{L_a}{R_a} \ll \frac{J}{B}$, and 2) if the tank time constant is very large compared with the complete motor time constant, i.e., if $RC \gg \frac{1}{\frac{B}{J} + \frac{k_m k_e}{J R_a}}$.

Finally, it is important to say that this procedure is valid because $\frac{B}{J} + \frac{k_m k_e}{J R_a} > 0$ and $\frac{R_a}{L_a} > 0$, i.e., the neglected poles are negative.

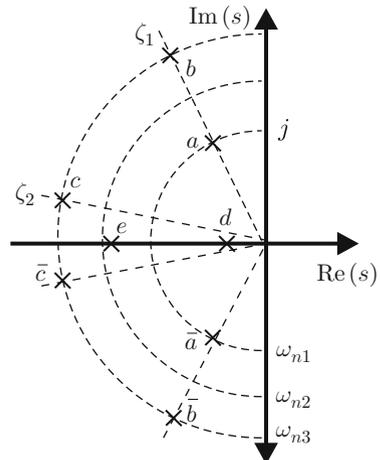
3.5.3 Approximating the Transient Response of Higher-Order Systems

The simplicity of first- and second-order systems allows us to compute their exact solutions. However, in the case of systems represented by differential equations of an order greater than 2, also called higher-order systems, the complexity of the problem prohibits us from obtaining their exact solutions for control purposes. The traditional treatment of higher-order systems in classical control has been to approximate their response as a kind of extrapolation of responses obtained in first- and second-order systems.

According to the studies presented in Sects. 3.1 and 3.3, the poles of first- and second-order systems can be located on the s complex plane, as depicted in Fig. 3.22. Then, the distance of a pole to the origin, $s = 0$, is representative of the response time: the larger this distance, the shorter the response time. The angle between the imaginary axis and the location vector of the pole is representative of how damped the response is: the closer this angle is to zero, the more oscillatory the response; conversely, the closer this angle is to 90° , the less oscillatory the response.

Fig. 3.22 Relative pole location on the s plane.

$\zeta_1 < \zeta_2$ and
 $\omega_{n1} < \omega_{n2} < \omega_{n3}$



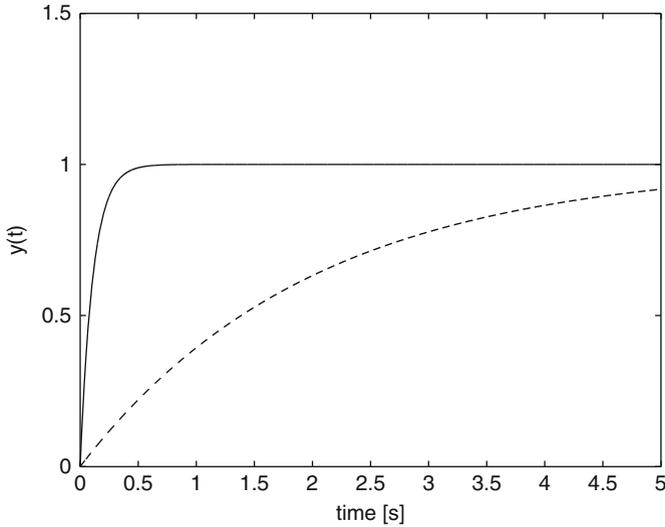


Fig. 3.23 Relative response of two systems with real poles (see Fig. 3.22). Dashed: pole at $s = d$. Continuous: pole at $s = e$

The responses of two first-order systems are compared in Fig. 3.23. The location of the respective real poles is depicted in Fig. 3.22. It is observed that a faster response is obtained, as the pole is located farther to the left of the origin.

The responses of three second-order systems are compared in Fig. 3.24. The location of the respective pair of complex conjugate poles is depicted in Fig. 3.22. It is observed that (i) a shorter response time is achieved as poles are located farther from the origin, and (ii) a less oscillatory response is achieved as the poles are located closer to the real axis.

The above results are applied to higher-order systems by considering the worst case, i.e., by observing the “slowest” pole (the closest pole to the origin) and the “least damped” pole (the closest pole to the imaginary axis).

3.6 The Case of Sinusoidal Excitations

Consider the n -order arbitrary transfer function:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{E(s)}{N(s)}, \tag{3.108}$$

$$N(s) = s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0,$$

$$E(s) = b_0 + b_1s + \dots + b_ms^m,$$

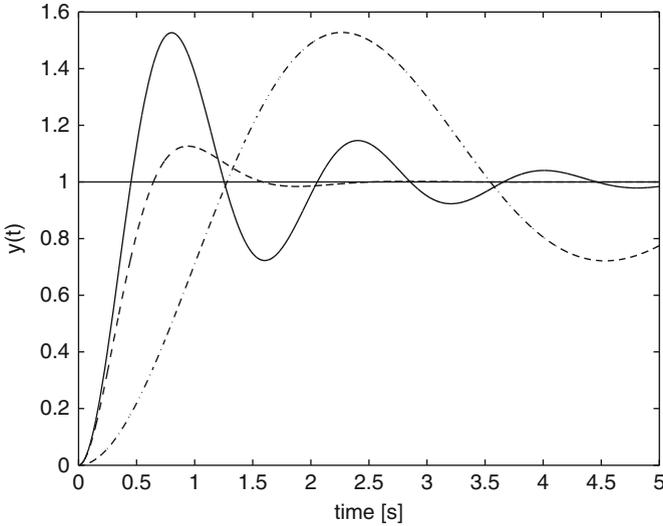


Fig. 3.24 Relative response of three systems with complex conjugate poles (see Fig. 3.22). Continuous: poles at $s = b$ and $s = \bar{b}$. Dash-dot: poles at $s = a$ and $s = \bar{a}$. Dashed: poles at $s = c$ and $s = \bar{c}$

where $n \geq m$ and $u(t) = A \sin(\omega t)$, i.e., $U(s) = A \frac{\omega}{s^2 + \omega^2}$. For the sake of simplicity, assume that $N(s)$ has no roots on the imaginary axis. Replacing $U(s)$ in (3.108) and using partial fraction expansion:

$$Y(s) = G(s)A \frac{\omega}{s^2 + \omega^2} = Y_n(s) + \frac{Cs + D}{s^2 + \omega^2}, \quad (3.109)$$

where $Y_n(s) = \mathcal{L}\{y_n(t)\}$ is the natural response, i.e., those terms arising from the roots of $N(s)$. Recall that $\lim_{t \rightarrow \infty} y_n(t) = 0$ if $G(s)$ is stable. As $N(s)$ has been assumed not to have any roots on the imaginary axis, then $G(s)$ has no poles at $s = \pm j\omega$. Multiplying both sides of (3.109) by the factor $s^2 + \omega^2$ and evaluating at $s = j\omega$, it is found:

$$\omega A G(j\omega) = j\omega C + D.$$

Equating real and imaginary parts yields:

$$C = A \operatorname{Im}(G(j\omega)), \quad D = \omega A \operatorname{Re}(G(j\omega)),$$

where $G(j\omega) = \operatorname{Re}(G(j\omega)) + j \operatorname{Im}(G(j\omega))$. Hence:

$$\frac{Cs + D}{s^2 + \omega^2} = A \operatorname{Im}(G(j\omega)) \frac{s}{s^2 + \omega^2} + A \operatorname{Re}(G(j\omega)) \frac{\omega}{s^2 + \omega^2}.$$

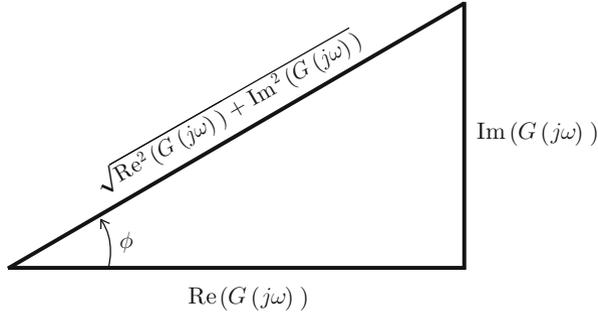


Fig. 3.25 Triangle defined by phase ϕ

Using the tabulated Laplace transforms:

$$\mathcal{L}\{\cos(\omega t)\} = \frac{s}{s^2 + \omega^2}, \quad \mathcal{L}\{\sin(\omega t)\} = \frac{\omega}{s^2 + \omega^2},$$

the following is found:

$$\mathcal{L}^{-1} \left\{ \frac{Cs + D}{s^2 + \omega^2} \right\} = A[\text{Im}(G(j\omega)) \cos(\omega t) + \text{Re}(G(j\omega)) \sin(\omega t)].$$

Using Fig. 3.25 and the trigonometric identity, $\sin(\alpha) \cos(\beta) + \cos(\alpha) \sin(\beta) = \sin(\alpha + \beta)$, the following is obtained:

$$\begin{aligned} \mathcal{L}^{-1} \left\{ \frac{Cs + D}{s^2 + \omega^2} \right\} &= A|G(j\omega)| [\sin(\phi) \cos(\omega t) + \cos(\phi) \sin(\omega t)], \quad (3.110) \\ &= B \sin(\omega t + \phi), \\ B &= A|G(j\omega)|, \\ \phi &= \arctan \left(\frac{\text{Im}(G(j\omega))}{\text{Re}(G(j\omega))} \right), \\ |G(j\omega)| &= \sqrt{\text{Re}^2(G(j\omega)) + \text{Im}^2(G(j\omega))}. \end{aligned}$$

Then, according to (3.109), the following is found:

$$y(t) = y_n(t) + y_f(t), \quad (3.111)$$

where the forced response is given as:

$$y_f(t) = B \sin(\omega t + \phi), \quad (3.112)$$

$$B = A|G(j\omega)|, \quad (3.113)$$

$$\phi = \arctan \left(\frac{\text{Im}(G(j\omega))}{\text{Re}(G(j\omega))} \right), \quad (3.114)$$

whereas $y_n(t) \rightarrow 0$ as $t \rightarrow \infty$ if $G(s)$ is stable. Hence, if the input $u(t)$ of an n -order linear system $G(s)$ is a sine function of time, then the output $y(t)$ is also, in a steady state, a sine function of time with the same frequency, but with an amplitude and a phase that are, in general, different from the amplitude and the phase of the signal at the input.

The ratio between the output and the input amplitudes at frequency ω , i.e., $\frac{B}{A} = |G(j\omega)|$, is known as the *magnitude of the transfer function* whereas the difference between the phase at the input and the phase at the output, ϕ , is known as the *phase of the transfer function*. Then, the following is concluded, which is an important result for the purposes of Chap. 6:

The frequency response is defined as the description of:

- How the ratio of the output amplitude and the input amplitude changes as the frequency of the signal at the input changes.
- How the difference between the phases of the signals at the input and the output changes as the frequency of the signal at the input changes.

Example 3.20 (Taken from [2], Ch. 4) Consider the spring-mass-damper system depicted in Fig. 3.18. Assume that there is no damper, i.e., $b = 0$, and the following external force $f = F_0 \sin(\omega t)$ is applied, where $\omega = \sqrt{\frac{K}{m}}$. It is desired to compute the mass position $x(t)$ if $x(0) = 0$ and $\dot{x}(0) = 0$. It is not necessary to compute the numerical value of constants appearing in $x(t)$.

Solution. The differential equation describing the situation in Fig. 3.18 when $b = 0$ is:

$$m\ddot{x} + Kx = f, \quad f = F_0 \sin(\omega t), \quad \omega = \sqrt{\frac{K}{m}}. \quad (3.115)$$

Applying the Laplace transform to (3.115) yields:

$$s^2 X(s) + \frac{K}{m} X(s) = \frac{1}{m} F(s).$$

Notice that this expression has the form:

$$s^2 X(s) + \omega_n^2 X(s) = \gamma \omega_n^2 F(s),$$

where $\omega_n = \sqrt{\frac{K}{m}} = \omega$, $\gamma = \frac{1}{K}$. As:

$$F(s) = \frac{F_0 \omega}{s^2 + \omega^2},$$

then:

$$X(s) = \frac{\gamma \omega^3 F_0}{(s^2 + \omega^2)^2}.$$

Using partial fraction expansion:

$$X(s) = \frac{As + B}{(s^2 + \omega^2)^2} + \frac{Cs + D}{s^2 + \omega^2}.$$

Thus, according to Sect. 3.4.4:

$$x(t) = \beta_1 t \sin(\omega t + \phi_1) + \beta_2 \sin(\omega t + \phi_2),$$

where $\beta_1, \beta_2, \phi_1, \phi_2$ are some real constants. The solution $x(t)$ is shown in Fig. 3.26 when $\omega = \sqrt{\frac{K}{m}} = 10[\text{rad/s}]$, $m = 1[\text{Kg}]$ and $F_0 = 1[\text{N}]$. This phenomenon is known as *resonance* and it means that the output reaches very large values despite the input being a bounded sine function of time. Notice that this may happen even if the characteristic polynomial has no roots with positive real parts nor located at $s = 0$. Thus, this is a different phenomenon to that described in Example 3.14. Also notice that resonance appears when a system is poorly damped ($\zeta \approx 0$) and the input is an oscillatory signal whose frequency is the same as (or very close to) the system natural frequency. As the mass position grows without a limit, resonance is a phenomenon that is considered to be dangerous.

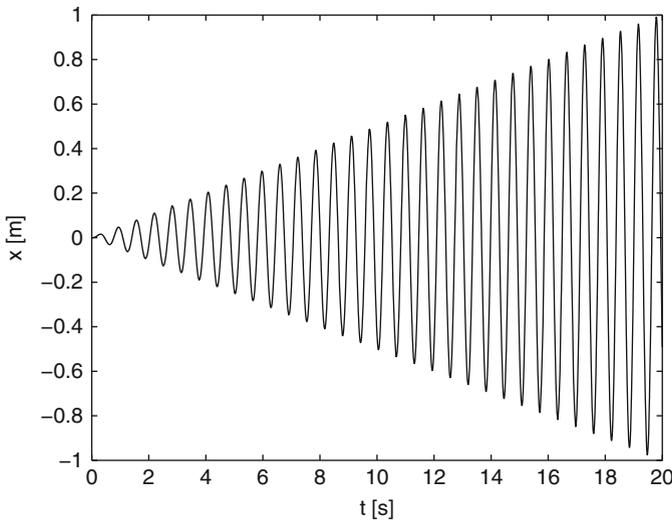


Fig. 3.26 Position in a mass-spring-damper system under resonance conditions

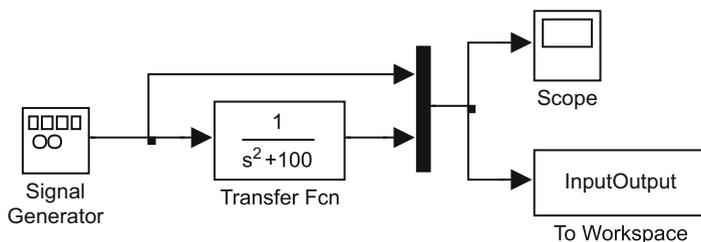


Fig. 3.27 MATLAB/Simulink simulation diagram for results in Fig. 3.26

Figure 3.26 has been obtained using the MATLAB/Simulink simulation diagram presented in Fig. 3.27. The *signal generator block* is programmed to have a sine wave form, unit amplitude and 10[rad/s] as frequency. Once the simulation stops, the following MATLAB code is executed in an m-file to draw Fig. 3.26:

```
nn=length(InputOutput(:,2));
n=nn-1;
Ts=20/n;
t=0:Ts:20;
plot(t,InputOutput(:,2),'k-');
axis([0 20 -1 1])
xlabel('t [s]')
ylabel('x [m]')
```

Example 3.21 Consider a linear system with transfer function:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{10}{s+5}, \quad y(0) = 0.$$

Suppose that $u(t) = A \sin(5t + 90^\circ)$, with $A = 1$. As $G(s)$ only has one pole at $s = -5$ and $\omega = 5$ in this case, according to (3.109), $Y_n(s) = \frac{h}{s+5}$ where:

$$h = (s+5) \frac{10}{s+5} \frac{\omega}{s^2 + \omega^2} \Big|_{s=-5} = \frac{10 \times 5}{5^2 + 5^2} = 1.$$

Hence, according to (3.111), (3.112), (3.113), (3.114):

$$y(t) = e^{-5t} + B \sin(5t + 90^\circ + \phi),$$

$$B = A|G(j\omega)|, \quad \phi = \arctan\left(\frac{\text{Im}(G(j\omega))}{\text{Re}(G(j\omega))}\right).$$

i.e.,:

$$\begin{aligned}
 G(j\omega) &= \frac{10}{j\omega + 5} \left(\frac{-j\omega + 5}{-j\omega + 5} \right), \\
 &= \frac{10(-j\omega + 5)}{25 + \omega^2}, \\
 &= \frac{10\sqrt{25 + \omega^2}}{\sqrt{25 + \omega^2}} \angle \arctan\left(\frac{-\omega}{5}\right), \\
 |G(j\omega)| &= \frac{10}{\sqrt{25 + \omega^2}}, \quad \phi = \arctan\left(\frac{-\omega}{5}\right).
 \end{aligned}$$

Evaluating at the applied frequency, $\omega = 5$:

$$|G(j\omega)| = \frac{10}{\sqrt{25 + 25}} = \frac{2}{\sqrt{2}} = \sqrt{2}, \quad \phi = -45^\circ.$$

Thus:

$$y(t) = e^{-5t} + \sqrt{2} \sin(5t + 45^\circ).$$

This response is depicted in Fig. 3.28. Notice that the difference between $y(t)$ and $y_f(t)$ at $t = 0.2 = 1/5$ [s] is equal to $1.38 - 1.015 = 0.3650 \approx e^{-5(0.2)}$, i.e., the natural response $y_n(t) = e^{-5t}$ evaluated at $t = 0.2$ [s]. This fact corroborates the above results.

Figure 3.28 has been obtained using the MATLAB/Simulink diagram presented in Fig. 3.29. The *sine wave block* is programmed to be a time-based block, with unit amplitude, zero bias, 5[rad/s] as the frequency, and 1.57[rad], i.e., 90° , as the phase. The *sine wave 1 block* is programmed to be a time-based block, with 1.4142 as amplitude, zero bias, 5[rad/s] as the frequency, and 1.57/2[rad], i.e., 45° , as the phase. Once the simulation stops, the following MATLAB code is executed in an m-file to draw Fig. 3.28:

```

nn=length(InputOutput(:,1));
n=nn-1;
Ts=2/n;
t=0:Ts:2;
plot(t,InputOutput(:,1),'k-.',t,InputOutput(:,2),
'k-');
hold on
plot(t,InputOutput(:,3),'k--');
axis([0 2 -1.5 1.5])
xlabel('time [s]')
hold off

```

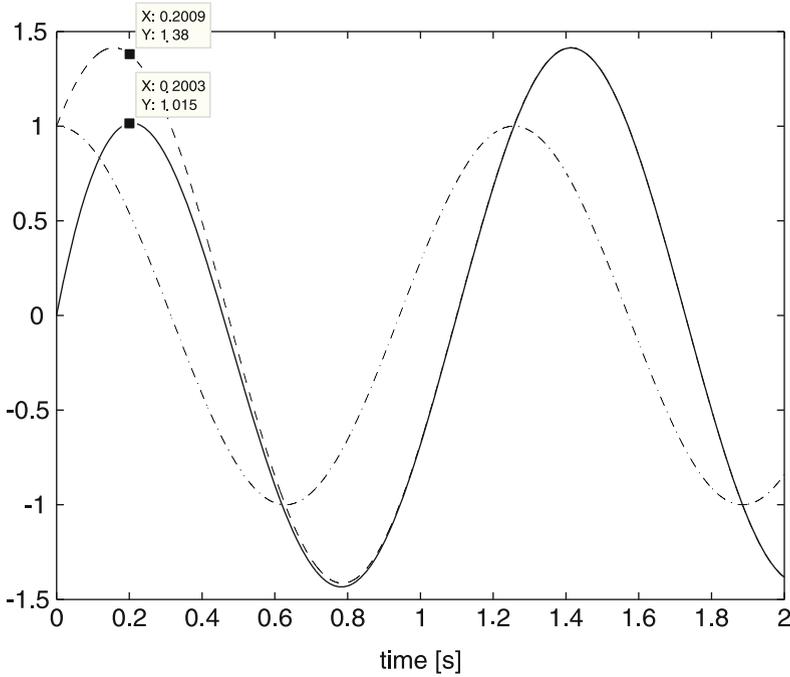


Fig. 3.28 Time response of a linear differential equation when excited with a sinusoidal function of time. Dash-dot: $u(t) = A \sin(5t + 90^\circ)$, $A = 1$. Dashed: $y_f(t) = \sqrt{2} \sin(5t + 45^\circ)$. Continuous: $y(t) = e^{-5t} + \sqrt{2} \sin(5t + 45^\circ)$

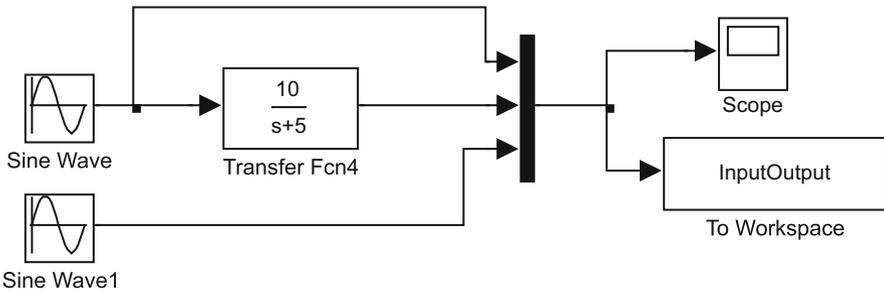


Fig. 3.29 MATLAB/Simulink diagram for the results in Fig. 3.28

3.7 The Superposition Principle

Every linear differential equation satisfies the superposition principle. Moreover, the fact that a differential equation satisfies the superposition principle is accepted as proof that the differential equation is linear. To simplify the exposition of ideas, the superposition principle is presented in the following only when all the initial conditions are zero.

Superposition Principle Consider the following n -order linear ordinary differential equation with constant coefficients:

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1\dot{y} + a_0y = b_0u + b_1\dot{u} + \cdots + b_mu^{(m)}, \quad (3.116)$$

where $n \geq m$. Assume that all of the initial conditions are zero. Also assume that $y_1(t)$ is the solution of (3.116) when $u_1(t)$ is applied at the input and $y_2(t)$ is the solution of (3.116) when $u_2(t)$ is applied at the input. Then, $\alpha_1y_1(t) + \alpha_2y_2(t)$, where α_1 and α_2 are arbitrary constants, is the solution of (3.116) when $\alpha_1u_1(t) + \alpha_2u_2(t)$ is applied at the input.

One way of verifying this result is presented in the following. As the differential equation in (3.116) is linear and all the initial conditions are zero, then it can be expressed in terms of the corresponding transfer function:

$$Y(s) = G(s)U(s). \quad (3.117)$$

This means that it can be written as follows:

$$Y_1(s) = G(s)U_1(s),$$

$$Y_2(s) = G(s)U_2(s).$$

Adding these expressions yields:

$$\begin{aligned} \alpha_1Y_1(s) + \alpha_2Y_2(s) &= \alpha_1G(s)U_1(s) + \alpha_2G(s)U_2(s), \\ &= G(s)(\alpha_1U_1(s) + \alpha_2U_2(s)). \end{aligned}$$

This is possible because α_1 and α_2 are constants. This expression corroborates that $\alpha_1y_1(t) + \alpha_2y_2(t)$ is the solution of (3.117); hence, the solution of (3.116) when $\alpha_1u_1(t) + \alpha_2u_2(t)$ is the input.

Example 3.22 Obtain the indicated voltage at impedance $Z_4(s)$ in Fig. 3.30a. This circuit can be simplified using the so-called *source transformation theorem*.

Theorem 3.1 (Source Transformation [5, 11], Chap. 9, pp. 61)

- When an impedance, $Z(s)$, and a voltage source, $V_{fv}(s)$, are series-connected between two terminals a and b , they can be replaced by a current source, $I_{fc}(s)$, parallel-connected to the same impedance, $Z(s)$. The magnitude of the current source is given as $I_{fc}(s) = V_{fv}(s)/Z(s)$.
- When an impedance, $Z(s)$, and a current source, $I_{fc}(s)$, are parallel-connected between two terminals a and b , they can be replaced by a voltage source, $V_{fv}(s)$, series-connected to the same impedance, $Z(s)$. The magnitude of the voltage source is given as $V_{fv}(s) = Z(s)I_{fc}(s)$.

Applying the first part of this theorem to voltage sources $V_1(s)$ and $V_2(s)$, circuit in Fig. 3.30b is obtained where:

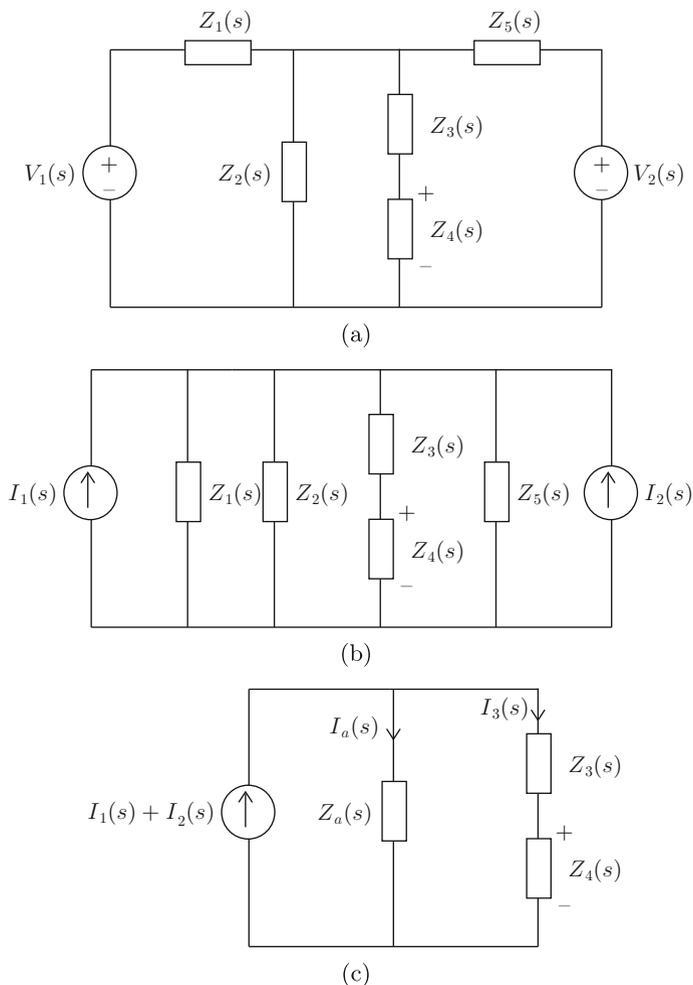


Fig. 3.30 Electric circuit studied in Example 3.22

$$I_1(s) = \frac{V_1(s)}{Z_1(s)}, \quad I_2(s) = \frac{V_2(s)}{Z_5(s)}. \quad (3.118)$$

It is clear that $Z_1(s)$, $Z_2(s)$ and $Z_5(s)$ are parallel-connected and its equivalent impedance is given as (see (2.69)):

$$Z_a(s) = \frac{1}{\frac{1}{Z_1(s)} + \frac{1}{Z_2(s)} + \frac{1}{Z_5(s)}}.$$

Using this and combining the current sources, the circuit in Fig. 3.30c is obtained. Now, another important result in network analysis is employed: the current divider.

Fact 3.1 (Current Divider [5], pp. 38) *When two parallel-connected impedances are parallel-connected to a current source, the current flowing through any of the impedances is given by the value of the current source multiplied by the opposite impedance to impedance where the current is to be computed and divided by the addition of the two impedances.*

Applying this result in addition to Ohm's Law to the circuit in Fig. 3.30c, the following expressions are found:

$$I_3(s) = \frac{Z_a(s)}{Z_a(s) + Z_3(s) + Z_4(s)} (I_1(s) + I_2(s)),$$

$$V_{z4}(s) = Z_4(s)I_3(s),$$

where $V_{z4}(s)$ is voltage at impedance $Z_4(s)$. Combining these expressions yields:

$$V_{z4}(s) = \frac{Z_a(s)Z_4(s)}{Z_a(s) + Z_3(s) + Z_4(s)} (I_1(s) + I_2(s)).$$

Finally, using (3.118), the following is found:

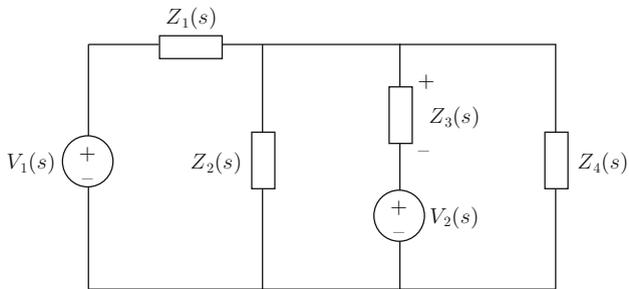
$$V_{z4}(s) = \frac{Z_a(s)Z_4(s)}{Z_a(s) + Z_3(s) + Z_4(s)} \left(\frac{V_1(s)}{Z_1(s)} + \frac{V_2(s)}{Z_5(s)} \right).$$

It is concluded that the voltage at $Z_4(s)$ can be obtained, as the addition of the voltages at $Z_4(s)$ due to each one of the voltage sources, i.e., $V_1(s)$ or $V_2(s)$, when the other source is put in short circuit (i.e., a zero value is assigned) and adding both results. This is exactly what the superposition principle establishes. It is interesting to say that the superposition principle has been established above in the present section, assuming that the different applied inputs are directly added, and, after that, they are applied to the system. However, this example shows that the superposition principle is valid in linear systems, even when the voltage sources, $V_1(s)$ and $V_2(s)$, cannot be added directly (see Fig. 3.30a).

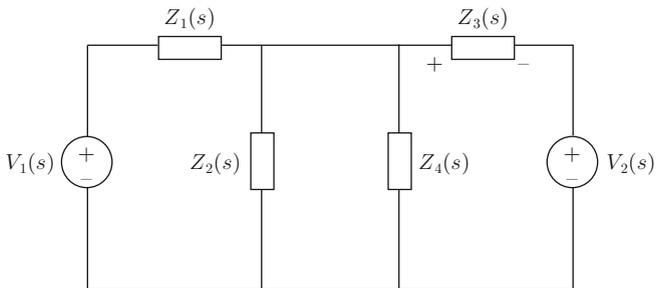
Example 3.2 Computing of the voltage at impedance $Z_3(s)$ in the circuit shown in Fig. 3.31a is required. First, this circuit is simplified to obtain circuit in Fig. 3.31b. Then, the source transformation theorem can be employed (see the previous example) to obtain the circuit in Fig. 3.31c, where:

$$I_1(s) = \frac{V_1(s)}{Z_1(s)}. \quad (3.119)$$

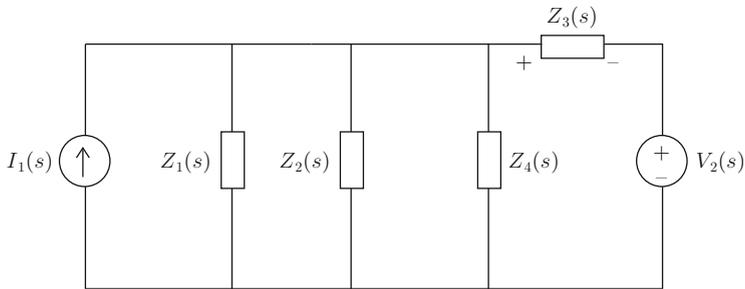
As impedances $Z_1(s)$, $Z_2(s)$, and $Z_4(s)$ are parallel-connected, the equivalent impedance is given as (see (2.69)):



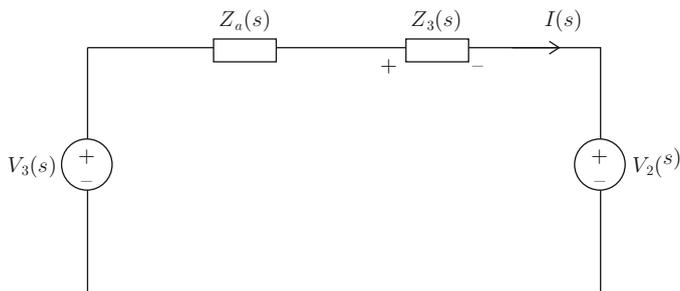
(a)



(b)



(c)



(d)

Fig. 3.31 Electric circuit studied in Example 3.2

$$Z_a(s) = \frac{1}{\frac{1}{Z_1(s)} + \frac{1}{Z_2(s)} + \frac{1}{Z_4(s)}}.$$

The equivalent impedance $Z_a(s)$ is parallel connected to the current source $I_1(s)$; hence, the second part of the source transformation theorem can be used (see the previous example) to obtain the circuit in Fig. 3.31d, where:

$$\begin{aligned} V_3(s) &= Z_a(s)I_1(s), \\ I(s) &= \frac{V_3(s) - V_2(s)}{Z_a(s) + Z_3(s)}, \\ V_{z3}(s) &= Z_3(s)I(s), \end{aligned}$$

with $V_{z3}(s)$ voltage at the impedance $Z_3(s)$. Combining these expressions and using (3.119) yields:

$$\begin{aligned} V_{z3}(s) &= \frac{Z_3(s)}{Z_a(s) + Z_3(s)}(V_3(s) - V_2(s)), \\ &= \frac{Z_3(s)}{Z_a(s) + Z_3(s)}(Z_a(s)I_1(s) - V_2(s)), \\ &= \frac{Z_3(s)}{Z_a(s) + Z_3(s)} \left(\frac{Z_a(s)}{Z_1(s)} V_1(s) - V_2(s) \right). \end{aligned}$$

It is concluded, again, that the voltage at the impedance $Z_3(s)$ can be obtained as the addition of the voltages in $Z_3(s)$ because of each one of the voltage sources, i.e., $V_1(s)$ or $V_2(s)$, when the other source is put in a short circuit (i.e., its value is set to zero) and adding the two results. This example also shows that the superposition principle is valid in linear systems, even when the voltage sources $V_1(s)$ y $V_2(s)$ cannot be added directly (see Fig. 3.31a).

3.8 Controlling First- and Second-Order Systems

The simplicity of the first- and second-order systems allows us to design feedback controllers whose effects can be explained relying on the exact solution of the closed-loop system. This is shown in the present section by applying this methodology to some physical systems.

3.8.1 Proportional Control of Velocity in a DC Motor

According to Chap. 10, the mathematical model of a permanent magnet brushed DC motor is given in (10.9), (10.10), which are repeated here for ease of reference:

$$\begin{aligned} L \frac{di}{dt} &= u - R i - n k_e \omega, \\ J \dot{\omega} &= -b \omega + n k_m i - T_p. \end{aligned} \quad (3.120)$$

The reader is encouraged to see Chap. 10 for an explanation of variables in this model. For the purposes of this example, it suffices to say that ω , u , i , T_p represent the load angular velocity, the applied voltage, the electric current through the armature circuit, and an external torque disturbance respectively, and all of the parameters are positive. If the motor is small, then it is widely accepted that $L \approx 0$ can be assumed, and the following is obtained:

$$0 = u - R i - n k_e \omega \quad \Rightarrow \quad i = \frac{u - n k_e \omega}{R}. \quad (3.121)$$

Replacing the electric current in (3.120) and rearranging:

$$J \dot{\omega} + \left(b + \frac{n^2 k_m k_e}{R} \right) \omega = \frac{n k_m}{R} u - T_p. \quad (3.122)$$

Consider the following proportional velocity controller:

$$u = k_p(\omega_d - \omega), \quad (3.123)$$

where k_p is a constant known as the proportional gain, and ω_d represents the desired velocity given as a step signal of magnitude A , i.e., $\omega_d(s) = \frac{A}{s}$. The closed-loop system is obtained by replacing u , given by the proportional controller, in the previous equation, i.e.,:

$$J \dot{\omega} + \left(b + \frac{n^2 k_m k_e}{R} \right) \omega = \frac{n k_m}{R} k_p (\omega_d - \omega) - T_p. \quad (3.124)$$

Rearranging terms yields:

$$\dot{\omega} + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \omega = \frac{n k_m k_p}{JR} \omega_d - \frac{1}{J} T_p.$$

This is a first-order system with two inputs. Then, we can apply superposition, i.e., define two signals ω_1 and ω_2 such that:

$$\begin{aligned}\omega(t) &= \omega_1(t) + \omega_2(t), \\ \dot{\omega}_1 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \omega_1 &= \frac{n k_m k_p}{JR} \omega_d, \\ \dot{\omega}_2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \omega_2 &= -\frac{1}{J} T_p.\end{aligned}$$

First, analyze the equation for ω_1 , applying the Laplace transform with zero initial conditions:

$$\begin{aligned}\omega_1(s) &= G_1(s)\omega_d(s), \\ G_1(s) &= \frac{\frac{n k_m k_p}{JR}}{s + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right)}.\end{aligned}\quad (3.125)$$

Notice that this transfer function is identical to:

$$\frac{b_2}{s + b_1}, \quad (3.126)$$

if we define:

$$b_1 = \frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \neq 0, \quad b_2 = \frac{n k_m k_p}{JR} \neq 0.$$

Notice that this expression has the same structure as the transfer function in (3.23). Hence, if the parameters and variables b_1 , b_2 , ω_1 and ω_d are suitably associated with those in (3.23), the solution $\omega_1(t)$ is the same as in (3.16), i.e.,:

$$\omega_1(t) = -\frac{b_2 A}{b_1} e^{-b_1 t} + \frac{b_2 A}{b_1} + \omega_1(0) e^{-b_1 t},$$

Stability is ensured if, and only if, the only pole of $G_1(s)$, located at $s = -b_1$, is negative, i.e., $-b_1 < 0$, which implies that:

$$k_p > -\frac{JR}{n k_m} \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right). \quad (3.127)$$

If this is the case, then the final value of the velocity can be computed using the final value theorem as:

$$\begin{aligned}\lim_{t \rightarrow \infty} \omega_1(t) &= \lim_{s \rightarrow 0} s \omega_1(s), \\ &= \lim_{s \rightarrow 0} s \frac{\frac{n k_m k_p}{JR}}{s + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right)} \frac{A}{s} = \frac{b_2 A}{b_1}.\end{aligned}$$

Recall that A represents the magnitude of the desired velocity. As it is desirable for the final value of the velocity to be close to A , it is also desirable for $\frac{b_2 A}{b_1}$ to have the same sign as A , i.e., $b_2/b_1 > 0$. As $b_1 > 0$, because of the stability requirement, then $b_2 > 0$ must also be true. This is ensured if $k_p > 0$, which is in complete agreement with (3.127). Hence, we can conclude that:

$$\lim_{t \rightarrow \infty} \omega_1(t) = \frac{b_2 A}{b_1} = \frac{\frac{n k_m k_p}{JR}}{\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR}} A < A, \quad (3.128)$$

which means that the final velocity remains below the desired velocity. This result may also be explained using everyday experience as follows. Assume for a moment that $\lim_{t \rightarrow \infty} \omega_1(t) = A$. Then, according to (3.123):

$$u = k_p(A - \omega) = 0,$$

the voltage applied at the motor terminals is zero. Hence, the motor tends to stop moving, i.e., ω_1 decreases to zero and cannot stay at the desired value A . This means that $\omega_1 < A$ will be obtained again. Hence, it is not possible to keep $\omega_1 = A$ forever, which justifies the result in (3.128).

The difference between the final velocity and A is known as the steady-state error e_{ss} . We remark that, according to (3.128), the steady-state error can be arbitrarily reduced by increasing $k_p > 0$, although a zero steady-state error will never be accomplished. On the other hand, if a specific steady-state error e_{ss} is required, the corresponding value for k_p can be computed as:

$$\begin{aligned} e_{ss} &= A - \frac{\frac{n k_m k_p}{JR}}{\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR}} A, \\ &= \frac{\frac{b}{J} + \frac{n^2 k_m k_e}{JR}}{\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR}} A, \\ k_p &= \left[\left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right) \frac{A}{e_{ss}} - \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right) \right] \frac{JR}{n k_m}. \end{aligned}$$

Notice that to ensure $k_p > 0$, the steady-state error is required to be a fraction of the desired velocity, i.e., $e_{ss} = \rho A$ with $0 < \rho < 1$.

Now, consider the expression for ω_2 . Using the Laplace transform:

$$\begin{aligned} \omega_2(s) &= G_2(s)T_p(s), \\ G_2(s) &= \frac{-\frac{1}{J}}{s + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right)}. \end{aligned}$$

Notice that $G_2(s)$ can be written as $G_2(s) = \frac{b_3}{s+b_1}$, with $b_3 = -\frac{1}{J}$, i.e., a transfer function such as that in (3.126). Thus, we may proceed as before to conclude that:

$$\lim_{t \rightarrow \infty} \omega_2(t) = \frac{-\frac{1}{J}}{\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR}} d, \quad T_p(s) = \frac{d}{s}.$$

The important observation in this expression is that the final value of ω_2 can be reduced by increasing $k_p > 0$. This is good news, as ω_2 is the velocity deviation due to external disturbance T_p ; hence, it is desirable to keep ω_2 close or equal to zero, if possible. In this case, the deviation can be rendered arbitrarily small by increasing $k_p > 0$, but a zero deviation will never be achieved with a proportional controller.

As the problem at hand is a first-order system, the transient response is completely described by the time constant:

$$\tau = \frac{1}{b_1} = \frac{1}{\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR}}.$$

Recall that a faster response is obtained as τ is shorter. This means that a faster response is obtained as $k_p > 0$ is chosen to be larger. Moreover, if all the system parameters are known and a desired time constant T is specified, then the required proportional gain can be exactly computed as:

$$k_p = \frac{JR}{n k_m} \left[\frac{1}{\tau} - \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right) \right].$$

This means that, to ensure $k_p > 0$, the closed-loop system is required to be faster than the open-loop system, i.e.:

$$\frac{1}{\tau} - \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right) > 0.$$

The velocity responses of a DC motor, when using two different proportional gains $k_p = 12$ and $k_p = 40$, are shown in Fig. 3.32. This was performed using the MATLAB/Simulink diagram shown in Fig. 3.33 to represent the closed-loop system in (3.124). It is assumed that $J = 1$, $n k_m / R = 1$ and $b + \frac{n^2 k_m k_e}{R} = 7$. These values were selected merely to obtain representative, clearly appreciable results for the proportional control of velocity. Block ω_d is a step applied at $t = 0$ with a zero initial value and 2 as a final value. The external disturbance T_p is a step applied at $t = 0.5$ with a zero initial value and 7 as a final value. Notice that the use of a larger k_p results in a faster response, i.e., $T_2 < T_1$, a smaller steady-state error, i.e., a larger steady-state response β , and a smaller deviation δ because of a constant external torque disturbance.

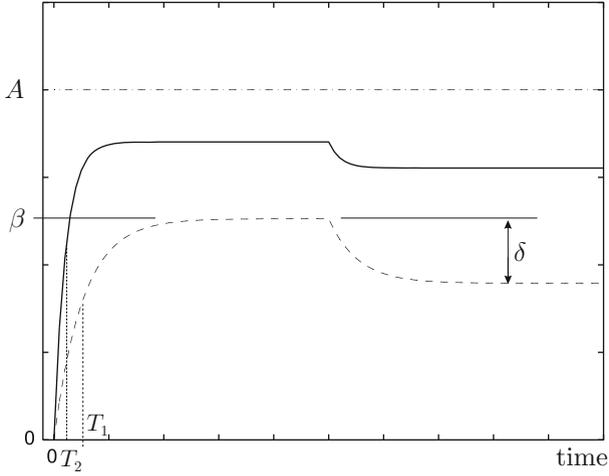


Fig. 3.32 Velocity response in the proportional control of velocity. Continuous: $k_p = 40$ is used. Dashed: $k_p = 12$ is used. Dash-dot: desired velocity

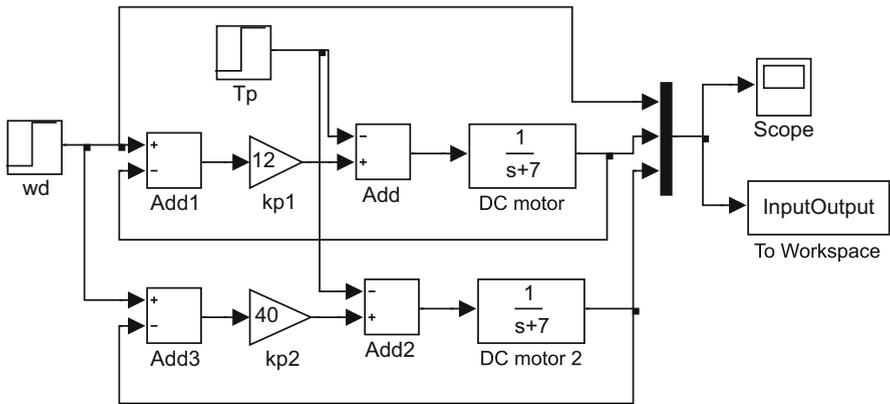


Fig. 3.33 MATLAB/Simulink diagram for the results in Fig. 3.32

The following is MATLAB code that is executed in an m-file to draw Fig. 3.32 after simulation in Fig. 3.33 stops:

```

nn=length(InputOutput(:,1));
n=nn-1;
Ts=1/n;
t=0:Ts:1;
plot(t,InputOutput(:,1),'k-.',t,InputOutput(:,3),
'k-');
hold on
    
```

```

plot(t, InputOutput(:,2), 'k--');
axis([-0.02 1 0 2.5])
xlabel('time [s]')
hold off

```

3.8.2 Proportional Position Control Plus Velocity Feedback for a DC Motor

Consider the motor model in (3.122), but, now, expressed in terms of position θ , i.e., $\dot{\theta} = \omega$:

$$J\ddot{\theta} + \left(b + \frac{n^2 k_m k_e}{R}\right) \dot{\theta} = \frac{n k_m}{R} u - T_p. \quad (3.129)$$

Assume that the voltage applied at the motor terminals is computed according to the following expression:

$$u = k_p(\theta_d - \theta) - k_v\dot{\theta},$$

where k_v is a constant known as the velocity gain, and θ_d represents the desired position, which is assumed to be given as a step signal with magnitude A , i.e., $\theta_d(s) = \frac{A}{s}$. This represents a controller known as *proportional position control plus velocity feedback*. The closed-loop equation is found by replacing u in (3.129) and rearranging terms, i.e.,:

$$\begin{aligned}
 J\ddot{\theta} + \left(b + \frac{n^2 k_m k_e}{R}\right) \dot{\theta} &= \frac{n k_m}{R} (k_p(\theta_d - \theta) - k_v\dot{\theta}) - T_p, & (3.130) \\
 \ddot{\theta} + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR}\right) \dot{\theta} + \frac{n k_m k_p}{JR} \theta &= \frac{n k_m k_p}{JR} \theta_d - \frac{1}{J} T_p,
 \end{aligned}$$

Using superposition and Laplace transform yields:

$$\begin{aligned}
 \theta(s) &= \theta_1(s) + \theta_2(s), \\
 \theta_1(s) &= G_1(s)\theta_d(s), \quad \theta_2(s) = G_2(s)T_p(s), \\
 G_1(s) &= \frac{\frac{n k_m k_p}{JR}}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR}\right) s + \frac{n k_m k_p}{JR}} \\
 G_2(s) &= \frac{-\frac{1}{J}}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR}\right) s + \frac{n k_m k_p}{JR}}.
 \end{aligned}$$

First, analyze the equation for θ_1 . By equating:

$$s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR} \right) s + \frac{n k_m k_p}{JR} = s^2 + 2\zeta\omega_n s + \omega_n^2,$$

it is possible to write:

$$G_1(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad (3.131)$$

$$\omega_n = \sqrt{\frac{n k_m k_p}{JR}} \quad \zeta = \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR} \right) \frac{1}{2\sqrt{\frac{n k_m k_p}{JR}}}.$$

Notice that, under these conditions, the inverse Laplace transform of $\theta_1(s) = G_1(s)\theta_d(s)$, is identical to the function presented in (3.57). Hence, stability is ensured if $k_p > 0$, to ensure that both ω_n and ζ are real numbers, and:

$$\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR} > 0,$$

to ensure that $\zeta > 0$. If this is the case, then we can use the final value theorem to compute the final position as:

$$\begin{aligned} \lim_{t \rightarrow \infty} \theta_1(t) &= \lim_{s \rightarrow 0} s\theta_1(s), \\ &= \lim_{s \rightarrow 0} s \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \frac{A}{s} = A. \end{aligned}$$

Thus, the motor position reaches the constant desired position in a steady state. On the other hand, according to (3.131), we conclude that:

- The closed-loop system is faster as $k_p > 0$ is chosen to be larger. This is because ω_n increases as $k_p > 0$ increases. However, as $k_p > 0$ increases the system becomes more oscillatory. This is because ζ decreases as k_p increases.
- The system response is less oscillatory as $k_v > 0$ is chosen to be larger. This is because ζ increases as k_v increases.

This means that the response can be rendered as fast and damped as desired merely by suitably selecting both k_p and k_v .

Now, consider the equation for ω_2 . Notice that this variable represents the position deviation produced by an external torque disturbance. Hence, it is desirable for ω_2 to be close or equal to zero, if possible. Notice that $G_2(s)$ is stable if is $G_1(s)$ stable, because both transfer functions have the same characteristic polynomial. Thus, we can compute the final value of ω_2 using the final value theorem when the disturbance is a step signal, i.e., $T_p(s) = \frac{d}{s}$:

$$\begin{aligned} \lim_{t \rightarrow \infty} \omega_2(t) &= \lim_{s \rightarrow 0} s \omega_2(s), \\ &= \lim_{s \rightarrow 0} s \frac{-\frac{1}{J} d}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_v}{JR} \right) s + \frac{n k_m k_p}{JR}}, \\ &= \frac{-\frac{1}{J} d}{\frac{n k_m k_p}{JR}}. \end{aligned}$$

This means that the steady-state deviation is not zero. However, the good news is that this deviation can be arbitrarily reduced by increasing k_p . We conclude that a faster closed-loop system also results in smaller position deviations because of constant external torque disturbances. This is corroborated by simulations shown in Fig. 3.34, which were performed using the MATLAB/Simulink diagram in Fig. 3.35. There, the closed-loop system in (3.130) is simulated using $J = 1$, $b + \frac{n^2 k_m k_e}{R} = 7$, and $n k_m / R = 70$. These values were employed merely because they allow representative, clearly appreciable, results to be obtained. The block *theta_d* is a step applied at $t = 0$ with a zero initial value and 2 as a final value. The external disturbance T_p is a step applied at $t = 0.5$ [s] with a zero initial value and 7×70 as a final value. Once simulation in Fig. 3.35 stops, the following MATLAB code is executed using an m-file to draw Fig. 3.34:

```
nn=length(InputOutput(:,1));
n=nn-1;
Ts=1/n;
```

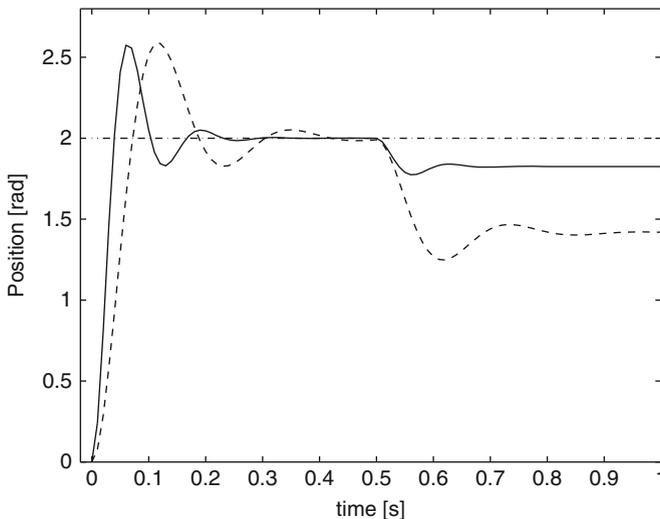


Fig. 3.34 Proportional control of the position with velocity feedback. Use of $k_p = 40$ and $k_v = 0.45$ (continuous) results in a faster response, but the same damping, than using $k_p = 12$ and $k_v = 0.2$ (dashed). See (3.132)

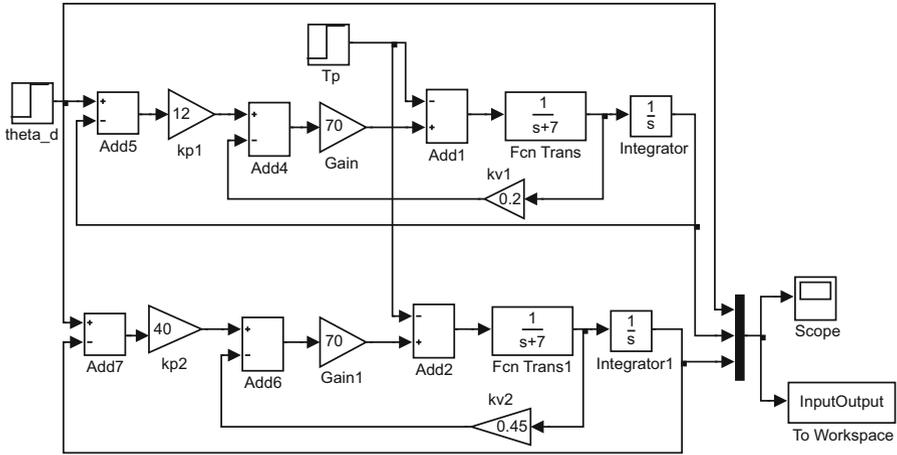


Fig. 3.35 MATLAB/Simulink diagram used to obtain results in Fig. 3.34

```

t=0:Ts:1;
plot(t, InputOutput(:,1), 'k-.', t, InputOutput(:,3),
'k-');
hold on
plot(t, InputOutput(:,2), 'k--');
axis([-0.02 1 0 2.75])
xlabel('time [s]')
ylabel('Position [rad]')
hold off

```

Notice that in the case where some t_r and M_p are specified, the use of expressions in (3.71) allow the required ω_n and ζ to be computed. These data and (3.131) allow the corresponding controller gains to be computed as:

$$k_p = \frac{JR}{n k_m} \omega_n^2, \quad k_v = \left[2\zeta \omega_n - \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right) \right] \frac{JR}{n k_m}. \quad (3.132)$$

Using this procedure together with $t_r = 0.039$ [s], $M_p = 29.9\%$, for the continuous line in Fig. 3.34, results in $k_p = 40$ and $k_v = 0.45$, whereas $t_r = 0.071$ [s], $M_p = 29.9\%$, for the dashed line, results in $k_p = 12$ and $k_v = 0.2$.

3.8.3 Proportional–Derivative Position Control of a DC Motor

Consider the motor model in (3.129) in a closed loop with the following proportional–derivative (PD) controller:

$$u = k_p(\theta_d - \theta) + k_d \frac{d}{dt}(\theta_d - \theta),$$

i.e.,:

$$\ddot{\theta} + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_d}{JR} \right) \dot{\theta} + \frac{n k_m k_p}{JR} \theta = \frac{n k_m k_p}{JR} \theta_d + \frac{n k_m k_d}{JR} \dot{\theta}_d - \frac{1}{J} T_p,$$

where k_d is known as the derivative gain. Using superposition and the Laplace transform yields:

$$\begin{aligned} \theta(s) &= \theta_1(s) + \theta_2(s), \\ \theta_1(s) &= G_1(s)\theta_d(s), \quad \theta_2(s) = G_2(s)T_p(s), \\ G_1(s) &= \frac{\frac{n k_m k_d}{JR}s + \frac{n k_m k_p}{JR}}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_d}{JR} \right) s + \frac{n k_m k_p}{JR}}, \\ G_2(s) &= \frac{-\frac{1}{J}}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_d}{JR} \right) s + \frac{n k_m k_p}{JR}}. \end{aligned}$$

Notice that except for the term $\frac{n k_m k_d}{JR}s$ in the numerator of $G_1(s)$ and the use of k_d instead of k_v in the denominator, we have exactly retrieved the same closed-loop equations as in the previous example. Thus, we can proceed exactly as in the previous example to conclude:

- Expressions in (3.131) stand again with k_d instead of k_v .
- Stability is ensured if $k_p > 0$ and:

$$\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_d}{JR} > 0.$$

- If any external disturbance is not present, then the position reaches the constant desired position in a steady state.
- The closed-loop system is faster as $k_p > 0$ is chosen to be larger. This is because ω_n increases as $k_p > 0$ increases. However, as $k_p > 0$ increases, the system becomes more oscillatory. This is because ζ decreases as k_p increases.
- The system response is less oscillatory as $k_d > 0$ becomes larger. This is because ζ increases as k_d increases.
- The response can be rendered as fast and damped as desired merely by suitably selecting both k_p and k_d .
- If a step external torque disturbance $T_p(s) = \frac{d}{s}$ is present, then a position deviation exists in a steady state and it is given as:

$$\lim_{t \rightarrow \infty} \omega_2(t) = \frac{-\frac{1}{J}}{\frac{n k_m k_p}{JR}} d.$$

On the other hand, to study the transient response notice that the transfer function $G_1(s)$ can be rewritten as:

$$\begin{aligned} G_1(s) &= \frac{\frac{k_d}{k_p} \omega_n^2 s + \omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2}, \\ &= \frac{k_d}{k_p} s \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} + \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2}, \end{aligned}$$

where:

$$\omega_n = \sqrt{\frac{n k_m k_p}{J R}}, \quad \zeta = \left(\frac{b}{J} + \frac{n^2 k_m k_e}{J R} + \frac{n k_m k_d}{J R} \right) \frac{1}{2\sqrt{\frac{n k_m k_p}{J R}}}, \quad (3.133)$$

have been used. Hence:

$$\theta_1(s) = \frac{k_d}{k_p} s \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} \theta_d(s) + \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} \theta_d(s). \quad (3.134)$$

As $\theta_d(s) = \frac{A}{s}$, notice that the inverse Laplace transform of the second term in this expression is the function presented in (3.57), whereas the inverse Laplace transform of the first term is the time derivative of the function in (3.57) multiplied by the factor $\frac{k_d}{k_p}$. This means that, although the shape of $\theta_1(t)$ is similar to that of the function in (3.57), there are some differences. For instance, formulas for the rise time and overshoot given in (3.71) are only approximate in the present control problem.

In the case in which some t_r and M_p are specified, use of the expressions in (3.71) allow only approximate values for the required ω_n and ζ to be computed. These data and (3.133) allow approximate values for the corresponding controller gains to be computed as:

$$k_p = \frac{J R}{n k_m} \omega_n^2, \quad k_d = \left[2\zeta \omega_n - \left(\frac{b}{J} + \frac{n^2 k_m k_e}{J R} \right) \right] \frac{J R}{n k_m}.$$

Some simulation results are presented in Fig. 3.36 when $J = 1$, $b + \frac{n^2 k_m k_e}{R} = 7$, and $n k_m / R = 70$. The corresponding MATLAB/Simulink diagram is shown in Fig. 3.37. The block theta_d is a step applied at $t = 0$ with a zero initial value and 2 as a final value. The external disturbance T_p is a step applied at $t = 0.5$ [s] with a zero initial value and 7×70 as a final value. When the simulation in Fig. 3.37 stops, the following MATLAB code is executed in an m-file to draw Fig. 3.36:

```
nn=length(InputOutput(:,1));
n=nn-1;
Ts=1/n;
t=0:Ts:1;
plot(t,InputOutput(:,2),'k-.',t,InputOutput(:,3),
```

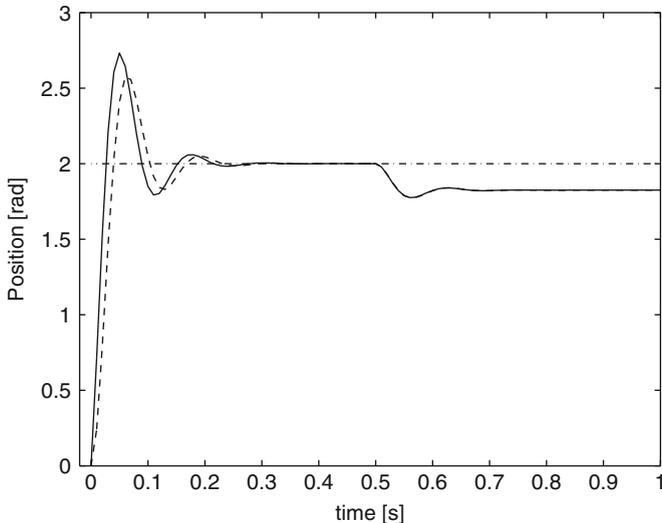


Fig. 3.36 Continuous: PD control law $u = k_p(\theta_d - \theta) + k_d \frac{d}{dt}(\theta_d - \theta)$, with $k_p = 40, k_d = 0.45$, is used. Dashed: control law $u = k_p(\theta_d - \theta) - k_v \dot{\theta}$, with $k_p = 40, k_v = 0.45$, is used. Dash-dot: desired position θ_d

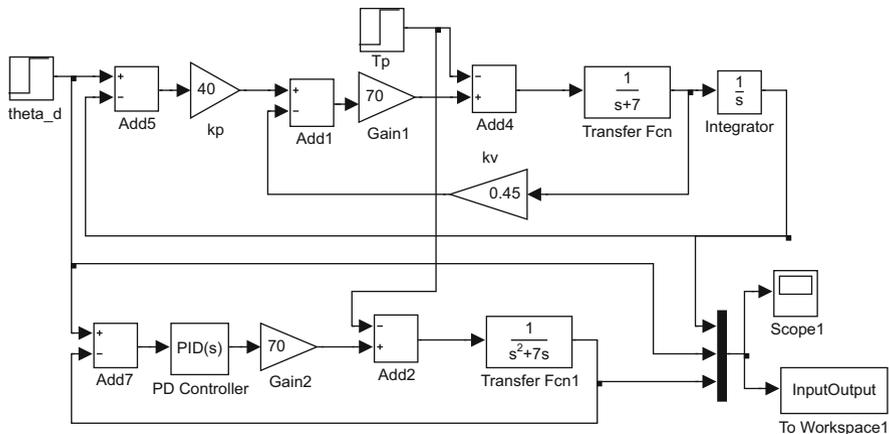


Fig. 3.37 MATLAB/Simulink diagram used to obtain results in Fig. 3.36

```
'k-');
hold on
plot(t, InputOutput(:, 1), 'k--');
axis([-0.02 1 0 3])
xlabel('time [s]')
ylabel('Position [rad]')
hold off
```

The dashed line represents the desired response, i.e., the inverse Laplace transform of:

$$\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}\theta_d(s), \quad (3.135)$$

which is identical to the response obtained with controller $u = k_p(\theta_d - \theta) - k_v\dot{\theta}$, introduced in the previous section, with $k_p = 40$, $k_v = 0.45$. The continuous line represents the inverse Laplace transform of the complete expression for $\theta_1(s)$ in (3.134), i.e., the response when a PD position controller is used with $k_p = 40$, $k_d = 0.45$. Notice that the continuous line in Fig. 3.36 has a smaller rise time and a larger overshoot than the dashed line. The reason for this is that the PD control law $u = k_p(\theta_d - \theta) + k_d\frac{d}{dt}(\theta_d - \theta)$ includes the term $k_d\frac{d\theta_d}{dt}$, which is not present in control law $u = k_p(\theta_d - \theta) - k_v\dot{\theta}$. Recall that θ_d is not a constant, but it is a step signal whose time derivative is very large at the time when such a step signal is applied. This contributes with a large voltage spike at this point in time, which results in a faster response, including a larger overshoot for the step response when a PD controller is employed.

Notice that a nonzero steady-state deviation results when a step external torque disturbance appears at $t = 0.5$ [s]. We stress that this deviation is the same in both responses, the continuous and the dashed line, because the same k_p is used.

3.8.4 Proportional–Integral Velocity Control of a DC Motor

Consider again the velocity control problem in a permanent magnet brushed DC motor. Now assume that the following proportional–integral (PI) controller is employed:

$$u = k_p(\omega_d - \omega) + k_i \int_0^t (\omega_d - \omega)dr,$$

where k_i is a constant known as the integral gain, whereas ω_d is a step signal standing for the desired velocity, i.e., $\omega_d(s) = \frac{A}{s}$. The closed-loop system is obtained by replacing u , given above, in the model (3.122), i.e.,:

$$J\dot{\omega} + \left(b + \frac{n^2 k_m k_e}{R}\right) \omega = \frac{n k_m}{R} \left(k_p(\omega_d - \omega) + k_i \int_0^t (\omega_d - \omega)dr\right) - T_p.$$

Differentiating once with respect to time and rearranging yields:

$$\ddot{\omega} + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR}\right) \dot{\omega} + \frac{n k_m k_i}{JR} \omega = \frac{n k_m k_p}{JR} \dot{\omega}_d + \frac{n k_m k_i}{JR} \omega_d - \frac{1}{J} \dot{T}_p. \quad (3.136)$$

This is a second-order system with two inputs. Then, superposition has to be used, i.e., to define two signals ω_1 and ω_2 such that:

$$\omega(t) = \omega_1(t) + \omega_2(t),$$

$$\ddot{\omega}_1 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \dot{\omega}_1 + \frac{n k_m k_i}{JR} \omega_1 = \frac{n k_m k_p}{JR} \dot{\omega}_d + \frac{n k_m k_i}{JR} \omega_d,$$

$$\ddot{\omega}_2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \dot{\omega}_2 + \frac{n k_m k_i}{JR} \omega_2 = -\frac{1}{J} \dot{T}_p.$$

First, analyze the equation for ω_1 . Applying the Laplace transform with zero initial conditions:

$$\begin{aligned} \omega_1(s) &= G_1(s)\omega_d(s). \\ G_1(s) &= \frac{\frac{n k_m k_p}{JR} s + \frac{n k_m k_i}{JR}}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) s + \frac{n k_m k_i}{JR}}. \end{aligned} \quad (3.137)$$

Since the characteristic polynomial is second-degree, we proceed to equate:

$$s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) s + \frac{n k_m k_i}{JR} = s^2 + 2\zeta\omega_n s + \omega_n^2,$$

to find:

$$\omega_n = \sqrt{\frac{n k_m k_i}{JR}} > 0, \quad \zeta = \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \frac{1}{2\sqrt{\frac{n k_m k_i}{JR}}}. \quad (3.138)$$

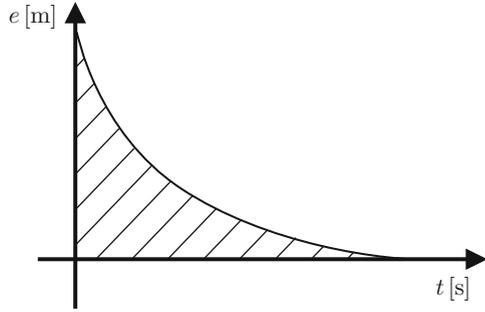
These expressions require $k_i > 0$ to ensure that ω_n and ζ are real numbers. On the other hand, $\zeta > 0$ is true if k_p is chosen to satisfy:

$$k_p > -\frac{JR}{n k_m} \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} \right).$$

Hence, $G_1(s)$ is a stable transfer function and, thus, the velocity final value can be computed using the final value theorem:

$$\begin{aligned} \lim_{t \rightarrow \infty} \omega_1(t) &= \lim_{s \rightarrow 0} s\omega_1(s), \\ &= \lim_{s \rightarrow 0} s \frac{\frac{n k_m k_p}{JR} s + \frac{n k_m k_i}{JR}}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) s + \frac{n k_m k_i}{JR}} \frac{A}{s}, \\ &= \frac{\frac{n k_m k_i}{JR}}{\frac{n k_m k_i}{JR}} A = A. \end{aligned}$$

Fig. 3.38 The integral $\int_0^t (\omega_d - \omega_1(r)) dr$ is represented by the shadowed area under function $e = \omega_d - \omega_1$. In this figure it is assumed that $\omega_d = A > \omega_1(0)$



It is concluded that the desired velocity is reached in a steady state, an important property of PI velocity control.

This result can be explained using everyday experience. As $\omega_1 = A$ is achieved in a steady state, the error signal $e = A - \omega(t)$ is expected to behave as depicted in Fig. 3.38. Recall that the integral operation $\int_0^t (\omega_d - \omega_1(r)) dr$ represents the area (positive) under the function shown in Fig. 3.38. This means that the above integral operation remains constant at a positive value when $\omega_1 = \omega_d = A$ (because the integrand becomes zero in these conditions). Then, according to $u = k_p(\omega_d - \omega) + k_i \int_0^t (\omega_d - \omega) dr$, a positive voltage given by $u = k_i \int_0^t (\omega_d - \omega) dr$ continues to be applied at the motor terminals. This allows the motor to stay rotating at the desired velocity, i.e., the condition $\omega_1 = \omega_d = A$ holds forever.

On the other hand, to study the transient response, notice that the transfer function in (3.137) can be rewritten as:

$$\begin{aligned} G_1(s) &= \frac{\frac{k_p}{k_i} \omega_n^2 s + \omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2}, \\ &= \frac{k_p}{k_i} s \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} + \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2}, \end{aligned}$$

where (3.138) has been used. Hence, according to (3.137):

$$\omega_1(s) = \frac{k_p}{k_i} s \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} \omega_d(s) + \frac{\omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} \omega_d(s).$$

As $\omega_d(s) = \frac{A}{s}$, notice that the inverse Laplace transform of the second term in this expression is the function presented in (3.57), whereas the inverse Laplace transform of the first term is the time derivative of the function in (3.57) multiplied by the factor $\frac{k_p}{k_i}$. This means that, although the shape of $\omega_1(t)$ is similar to that of the function in (3.57), there are some differences. For instance, the rise time and overshoot formulas given in (3.71) are only approximate in the present control problem. This is a similar situation to that found in the study of the transient

response of PD position control of a DC motor. Despite this, the expressions in (3.138) are still valid and they can be used to conclude the following:

- A faster response is achieved as $k_i > 0$ is chosen to be larger. This is because ω_n increases as $k_i > 0$ is increased. However, a more oscillatory response is obtained as $k_i > 0$ is chosen to be larger. This is because ζ decreases as k_i increases.
- The oscillation can be reduced by increasing $k_p > 0$. This is because ζ increases as $k_p > 0$ increases.

These observations can be used as a tuning procedure to select both k_p and k_i for a PI velocity controller. Note that this is a trial and error procedure.

Finally, consider the expression for ω_2 :

$$\ddot{\omega}_2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) \dot{\omega}_2 + \frac{n k_m k_i}{JR} \omega_2 = -\frac{1}{J} \dot{T}_p.$$

Using the Laplace transform yields:

$$\omega_2(s) = \frac{-\frac{1}{J}s}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) s + \frac{n k_m k_i}{JR}} T_p(s).$$

Then, the final value of ω_2 can be computed as:

$$\begin{aligned} \lim_{t \rightarrow \infty} \omega_2(t) &= \lim_{s \rightarrow 0} s \omega_2(s), \\ &= \lim_{s \rightarrow 0} s \frac{-\frac{1}{J}s}{s^2 + \left(\frac{b}{J} + \frac{n^2 k_m k_e}{JR} + \frac{n k_m k_p}{JR} \right) s + \frac{n k_m k_i}{JR}} \frac{d}{s} = 0, \end{aligned}$$

when the external torque disturbance is a step signal, i.e., $T_p(s) = \frac{d}{s}$. Recall that it has already been ensured that the characteristic polynomial in this expression has two poles with a negative real part. Hence, the steady-state velocity deviation due to a constant external torque disturbance is zero, which is very good news.

The closed-loop response of velocity under PI control is shown in Fig. 3.39 where $J = 1$, $b + \frac{n^2 k_m k_e}{R} = 7$, and $n k_m / R = 70$ are used. These values are employed merely because they allow representative, clear, responses to be obtained. The corresponding MATLAB/Simulink diagram is shown in Fig. 3.40. The block ω_d is a step applied at $t = 0$ with a zero initial value and 2 as a final value. The external disturbance T_p is a step applied at $t = 0.5$ [s] with a zero initial value and 0.4×70 as a final value. Once the simulation stops, the following MATLAB code is executed in an m-file to draw Fig. 3.39:

```
nn=length(InputOutput(:,1));
n=nn-1;
Ts=1/n;
```

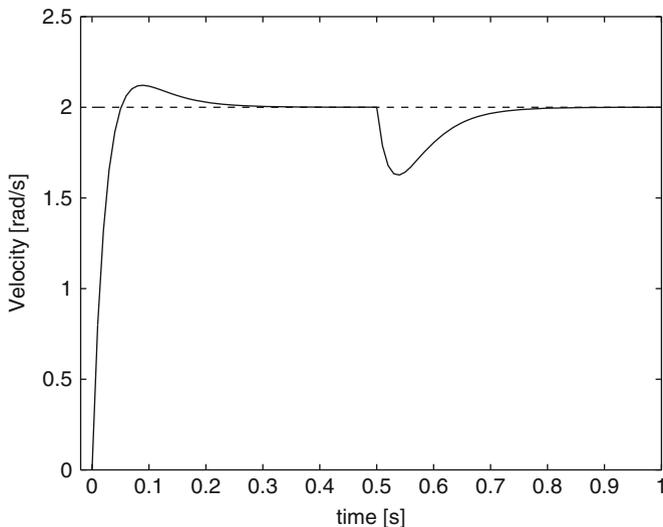


Fig. 3.39 PI velocity control when a constant external disturbance appears at $t = 0.5[s]$. The PI controller gains are $k_p = 0.7$ and $k_i = 10$. Continuous: velocity $\omega(t)$. Dashed: desired velocity ω_d

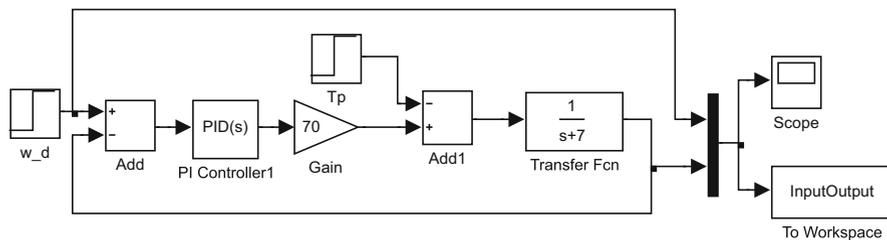


Fig. 3.40 MATLAB/Simulink diagram to obtain results in Fig. 3.39

```

t=0:Ts:1;
plot(t,InputOutput(:,1),'k--',t,InputOutput(:,2),
'k-');
axis([-0.02 1 0 2.5])
xlabel('time [s]')
ylabel('Velocity [rad/s]')
    
```

Notice that the velocity reaches its desired value before and after a constant external torque disturbance appears at $t = 0.5[s]$.

3.8.5 Proportional, PI, and PID Control of First-Order Systems

Consider the water level system studied in Example 3.2 and depicted in Fig. 3.3. Suppose that a water pump is employed to produce a water flow q_i that is proportional to the voltage u applied at motor terminals actuating on the pump, i.e.,:

$$q_i = k_1 u, \quad (3.139)$$

where k_1 is a positive constant. Combining (3.139) and (3.28) it is obtained:

$$\frac{dh}{dt} + ah = kk_1 u.$$

Notice that this differential equation can also be represented by a transfer function with the form:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_2}{s + b_1}. \quad (3.140)$$

Hence, identical conclusions are obtained as those for the proportional control of velocity, if a proportional level controller is used, or those for a proportional–integral velocity controller, if a PI level controller is used. Moreover, identical conclusions are valid for any first order plant with transfer function having the same structure as in (3.140).

On the other hand, first order systems represented by (3.140) do not need use of a controller with a derivative action. This can be explained as follows. Suppose that the following proportional–integral–derivative (PID) controller:

$$u = k_p e + k_d \frac{de}{dt} + k_i \int_0^t e(r) dr, \quad e = y_d - y$$

is used for plant in (3.140), i.e.,:

$$\begin{aligned} \dot{y} + b_1 y &= b_2 \left(k_p e + k_d \frac{de}{dt} + k_i \int_0^t e(r) dr \right), \\ (1 + b_2 k_d) \dot{y} + (b_1 + b_2 k_p) y &= b_2 k_p y_d + b_2 k_d \dot{y}_d + b_2 k_i \int_0^t e(r) dr, \end{aligned}$$

and differentiating once with respect to time:

$$(1 + b_2 k_d) \ddot{y} + (b_1 + b_2 k_p) \dot{y} + b_2 k_i y = b_2 k_d \ddot{y}_d + b_2 k_p \dot{y}_d + b_2 k_i y_d.$$

The corresponding closed-loop transfer function is:

$$H(s) = \frac{Y(s)}{Y_d(s)} = \frac{b_2k_d s^2 + b_2k_p s + b_2k_i}{(1 + b_2k_d)s^2 + (b_1 + b_2k_p)s + b_2k_i}. \quad (3.141)$$

Notice that the characteristic polynomial is second degree. Thus, only two free parameters are necessary to assign the corresponding two roots at any location on the s complex plane, i.e., to achieve any desired transient response. However, three free parameters appear in this characteristic polynomial, i.e., k_p , k_d and k_i . Hence, one of them is not necessary, i.e., k_d .

On the other hand, the transfer function in (3.141) has polynomials at the numerator and the denominator having the same degree, i.e., it has the same number of poles and zeros. It has been shown in Examples 3.6 and 3.7 that time response of a system whose transfer function has the same number of poles and zeros is discontinuous when a step input is applied. It is important to remark that such a response is not possible, for instance, in a velocity control system because the velocity cannot be discontinuous. This is due to the fact that such a response would require a very large voltage to be applied at the motor terminals during a very short time interval. Mathematically, this is explained by the time derivative of the desired value y_d (a discontinuous, step signal) which is introduced by the derivative action of a controller. However, such large voltage spikes are not possible in practice. Thus, an unpredictable response is expected. Moreover, in the case of a velocity control system, the derivative action on velocity (equivalent to acceleration measurements, a very noisy signal) results in a strong amplification of noise, as velocity measurements are recognized to be noisy signals.

Thus, it is concluded that any performance improvement is not expected when using a PID controller with respect to performance achieved with a PI controller for first-order plants.

3.9 Case Study: A DC-to-DC High-Frequency Series Resonant Power Electronic Converter

In Sect. 2.5, Chap. 2 the mathematical model of a DC-to-DC high-frequency series resonant power electronic converter was obtained. This model is given in (2.142)–(2.144) and it is rewritten here for ease of reference:

$$L \frac{di}{dt} = -v - v_0 \text{sign}(i) + E(t), \quad (3.142)$$

$$C \frac{dv}{dt} = i, \quad (3.143)$$

$$C_0 \frac{dv_0}{dt} = \text{abs}(i) - \frac{v_0}{R}, \quad (3.144)$$

where:

$$\text{sign}(i) = \begin{cases} +1, & i > 0 \\ -1, & i < 0 \end{cases}, \quad (3.145)$$

whereas $\text{abs}(i)$ stands for the absolute value of i . The operation of this circuit is studied in the present section through the solution of the mathematical model in (3.142)–(3.144). Performing a coordinate change for the variables involved is very useful. Hence, the following variables are defined (see Sect. 2.5, Chap. 2, for an explanation of the meaning of the parameters involved):

$$z_1 = \frac{i}{E} \sqrt{\frac{L}{C}}, \quad z_2 = \frac{v}{E}, \quad z_3 = \frac{v_0}{E}, \quad \tau = \frac{t}{\sqrt{LC}}. \quad (3.146)$$

Replacing this in (3.142), (3.143), (3.144) yields:

$$\begin{aligned} L \frac{d \left(E \sqrt{\frac{C}{L}} z_1 \right)}{d \left(\tau \sqrt{LC} \right)} &= -E z_2 - E z_3 \text{sign}(z_1) + E(t), \\ C \frac{d(E z_2)}{d \left(\tau \sqrt{LC} \right)} &= z_1 E \sqrt{\frac{C}{L}}, \\ C_0 \frac{d(E z_3)}{d \left(\tau \sqrt{LC} \right)} &= \text{abs} \left(E \sqrt{\frac{C}{L}} z_1 \right) - \frac{E z_3}{R}. \end{aligned}$$

After some simplifications, the following is found:

$$\dot{z}_1 = -z_2 - z_3 \text{sign}(z_1) + u, \quad (3.147)$$

$$\dot{z}_2 = z_1, \quad (3.148)$$

$$\alpha \dot{z}_3 = \text{abs}(z_1) - \frac{z_3}{Q}, \quad (3.149)$$

where the dot “.” represents the differentiation with regard to the normalized time τ whereas $\alpha = C_0/C$, $Q = R\sqrt{C/L}$ and u is a variable taking only the two possible values $+1$ (when $E(t) = +E$) and -1 (when $E(t) = -E$). The output capacitor C_0 is commonly chosen to be very large compared with the value of the resonant capacitor C , because this ensures that z_3 , i.e., v_0 , remains approximately constant during several cycles of the resonant current z_1 , i.e., i . Hence, under this assumption (a constant z_3), (3.149) can be neglected and (3.147), (3.148), suffice to represent the evolution of the resonant variables z_1 and z_2 , at least for several oscillations.

For operation at resonance, the transistors Q_1 , Q_3 are turned on when $z_1 > 0$, i.e., when $i > 0$; hence, $u = +1$ because $E(t) = +E$ in this case, according to

Figs. 2.34a, 2.35 and 2.36. On the other hand, the transistors Q_2, Q_4 are turned on when $z_1 < 0$, i.e., when $i < 0$, and hence $u = -1$ because $E(t) = -E$ in this case. One way of achieving this is by designing $u = \text{sign}(z_1)$, where $\text{sign}(z_1) = +1$ if $z_1 > 0$ and $\text{sign}(z_1) = -1$ if $z_1 < 0$. This means that, according to (3.147) and (3.148), the evolution of the resonant variables can be described as:

$$\dot{z}_1 = -z_2 + v_e, \quad (3.150)$$

$$\dot{z}_2 = z_1, \quad (3.151)$$

$$v_e = \begin{cases} 1 - z_3, & Q_1, Q_3 \text{ turned-on} \\ -(1 - z_3), & Q_2, Q_4 \text{ turned-on} \end{cases},$$

where z_3 is assumed to be constant. Using the variable change $z_4 = z_2 - v_e$ in the previous equations yields $\dot{z}_4 = \dot{z}_2 = z_1$, because a constant v_e implies that $\dot{v}_e = 0$, and $\ddot{z}_4 = \dot{z}_1 = -z_2 + v_e = -z_4$, i.e.,:

$$\ddot{z}_4 + z_4 = 0. \quad (3.152)$$

This is a second-order linear ordinary differential equation with constant coefficients such as that defined in (3.53) with $\zeta = 0$ and $\omega_n = 1$. As the excitation, or input, of this differential equation is zero, then $A = 0$ and the complete solution is given by the natural response alone. Hence, according to Sect. 3.3, i.e., expressions in (3.62) and (3.63), the solution of (3.152) is given as:

$$\begin{aligned} z_4(\tau) = p(\tau) &= \mathcal{L}^{-1} \left\{ \frac{z_{40}(s + 2\zeta\omega_n) + \dot{z}_{40}}{s^2 + 2\zeta\omega_n s + \omega_n^2} \right\} = \mathcal{L}^{-1} \left\{ \frac{z_{40}s + \dot{z}_{40}}{s^2 + 1} \right\}, \\ &= \mathcal{L}^{-1} \left\{ \frac{z_{40}s}{s^2 + 1} + \frac{\dot{z}_{40}}{s^2 + 1} \right\} = z_{40} \cos(\tau) + \dot{z}_{40} \sin(\tau), \end{aligned} \quad (3.153)$$

where the following Laplace transform pairs have been employed [4], Ch. 32:

$$\mathcal{L}^{-1} \left\{ \frac{s}{s^2 + a^2} \right\} = \cos(a\tau),$$

$$\mathcal{L}^{-1} \left\{ \frac{a}{s^2 + a^2} \right\} = \sin(a\tau),$$

whereas $z_{40} = z_4(0) = z_2(0) - v_e$ and $\dot{z}_{40} = \dot{z}_4(0) = z_1(0)$. Differentiating (3.153) once with respect to τ yields:

$$\dot{z}_4(\tau) = -z_{40} \sin(\tau) + \dot{z}_{40} \cos(\tau). \quad (3.154)$$

Then, using (3.153) and (3.154), it is easy to verify that:

$$z_4^2(\tau) + \dot{z}_4^2(\tau) = z_{40}^2 + \dot{z}_{40}^2. \quad (3.155)$$

If the solution of (3.152) is plotted on the phase plane $\dot{z}_4 - z_4$, then a circle is obtained whose center is located at origin $(z_4, \dot{z}_4) = (0, 0)$, and the radius is determined by the initial conditions, i.e., $\sqrt{z_{40}^2 + \dot{z}_{40}^2}$. Moreover, according to $z_4 = z_2 - v_e$ and $\dot{z}_4 = z_1$, it is concluded that the solution of (3.150), (3.151), when plotted on the phase plane $z_1 - z_2$, is a circle centered at $(z_2, z_1) = (v_e, 0)$, with a radius $\sqrt{(z_{20} - v_e)^2 + z_{10}^2}$ where $z_{20} = z_2(0)$ and $z_{10} = z_1(0)$. It is important to say that, according to (3.150) and (3.151), the values of z_1 and z_2 cannot be discontinuous because that would require infinite values for \dot{z}_1 and \dot{z}_2 respectively, which is not possible as the right-hand sides of (3.150) and (3.151) cannot take infinite values. This means that, when combining the solutions in both regions, i.e., when $z_1 > 0$ and when $z_1 < 0$, the initial conditions on each region must be equal to the final values reached in the previous region. Hence, Fig. 3.41 is obtained. The closed trajectory shown in this figure indicates that the resonant variables z_2, z_1 , i.e., v and i , have sustained oscillations. It should be noted that this closed trajectory is clockwise-oriented because, according to (3.151), z_2 grows when $z_1 > 0$, i.e., if $\dot{z}_2 > 0$ then z_2 grows, and z_2 decreases when $z_1 < 0$. Notice that v_e can only take two different values, $1 - z_3$ when $z_1 > 0$, and $-(1 - z_3)$ when $z_1 < 0$. Thus, the only way in which the situation represented in Fig. 3.41 can be obtained is that $z_3 = 1$, i.e., that $v_e = 0$, in both regions. This means that the output voltage is equal to the power supply voltage, i.e., $v_0 = E$.

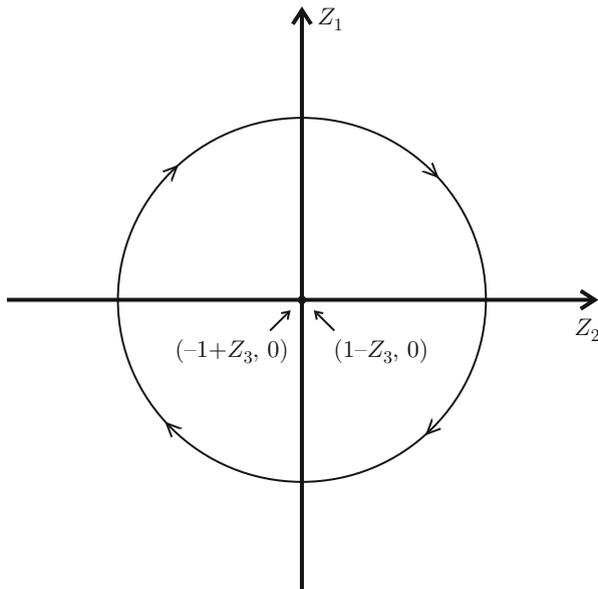


Fig. 3.41 Evolution of the resonant variables z_2 and z_1 when $u = \text{sign}(z_1)$

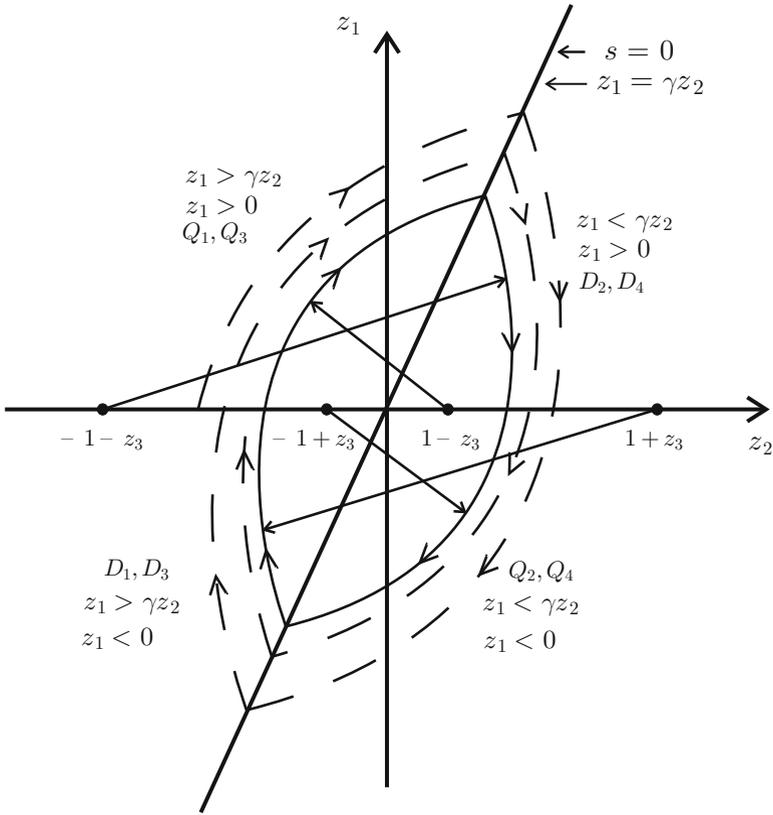


Fig. 3.42 Evolution of the resonant variables z_2 y z_1 when $u = \text{sign}(s)$ with $s = z_1 - \gamma z_2$

An important application of DC-to-DC high-frequency series resonant power electronic converters is to deliver output voltages that are smaller than the power supply voltage, i.e., $v_0 < E$ or $z_3 < 1$. Define a function $s = z_1 - \gamma z_2$, with γ a positive constant, and assign $u = \text{sign}(s)$. This defines the four regions shown in Fig. 3.42. Notice that the transistors Q_1, Q_3 are turned on when $u = +1$, i.e., above the straight line $z_1 = \gamma z_2$, but the electric current flows through them only when $z_1 > 0$ as, even when Q_1, Q_3 are turned on, the electric current flows through the diodes D_1, D_3 when $z_1 < 0$ (the electric current only flows in one direction through any of the transistors Q_1, Q_2, Q_3, Q_4). On the other hand, the transistors Q_2, Q_4 are turned on when $u = -1$, i.e., below the straight line $z_1 = \gamma z_2$, but the electric current flows through them only when $z_1 < 0$, as, despite Q_2, Q_4 being turned on, the electric current flows through diodes D_2, D_4 when $z_1 > 0$.

The resonant variables still evolve according to (3.150) and (3.151), but now v_e takes the following constant values:

$$v_e = \begin{cases} 1 - z_3, & Q_1, Q_3 \text{ in saturation, } (z_1 > \gamma z_2, z_1 > 0) \\ -(1 - z_3), & Q_2, Q_4 \text{ in saturation, } (z_1 < \gamma z_2, z_1 < 0) \\ 1 + z_3, & D_1, D_3 \text{ conduct, } (z_1 > \gamma z_2, z_1 < 0) \\ -(1 + z_3), & D_2, D_4 \text{ conduct, } (z_1 < \gamma z_2, z_1 > 0) \end{cases}$$

This means that, again, and according to arguments between (3.150) and the paragraph after (3.155), the resonant variables describe circles on the $z_2 - z_1$ plane. Radii of these circles are determined by the initial conditions in each region, which are equal to the final values in the previous region. The center of these circles is located at the following points, depending on the operation region:

$$\begin{aligned} (z_2, z_1) &= (1 - z_3, 0), & z_1 > \gamma z_2, z_1 > 0, \\ (z_2, z_1) &= (-1 + z_3, 0), & z_1 < \gamma z_2, z_1 < 0, \\ (z_2, z_1) &= (1 + z_3, 0), & z_1 > \gamma z_2, z_1 < 0, \\ (z_2, z_1) &= (-1 - z_3, 0), & z_1 < \gamma z_2, z_1 > 0. \end{aligned}$$

Figure 3.42 depicts the corresponding situation when $\gamma = 3$. Notice that, again, the resonant variables z_2, z_1 , i.e., v and i , describe sustained oscillations. To obtain the closed trajectory shown in Fig. 3.42, we may proceed as follows. Propose any $z_3 < 1$ and any point in the $z_2 - z_1$ plane as an initial condition. Using a pair of compasses, draw circles centered at the above indicated points according to the respective operation region. According to the previous discussion and (3.151), these circles are clockwise-oriented. The reader will realize that an helix appears similar to the dashed line in Fig. 3.42, which finally ends on a closed trajectory. The fact that this closed trajectory is reached from any initial point on the $z_2 - z_1$ plane is indicative that it is stable or attractive. It is possible to proceed analogously using some $z_3 > 1$ to verify that a closed trajectory is also reached in such a case. This is indicative that in a series resonant converter is not possible to deliver, in a steady state, an output voltage v_0 larger than the power supply voltage E , i.e., $z_3 > 1$ is not possible in a steady state. If an output voltage $v_0 > E$ is desired, then the parallel resonant converter shown in Fig. 2.34b is a suitable alternative.

On the other hand, arcs with a center at $(z_2, z_1) = (1 - z_3, 0)$ and $(z_2, z_1) = (-1 - z_3, 0)$, shown in Fig. 3.42, are drawn one after the other. Then, if both of them were centered at origin $(z_2, z_1) = (0, 0)$, they would together span a 180° angle. However, using basic geometry in Fig. 3.42 it is possible to conclude that arcs with the center at $(z_2, z_1) = (1 - z_3, 0)$ and $(z_2, z_1) = (-1 - z_3, 0)$ together describe an angle smaller than 180° despite completing a semicircle on both resonant variables z_2, z_1 . Hence, the four arcs composing the whole closed trajectory in Fig. 3.42 together describe an angle smaller than 360° despite a complete cycle being described on both resonant variables. Notice that each one of these arcs is covered at a frequency that is still $\omega_n = 1$, but, according to the above description, less time is now required to complete an oscillation as the four arcs together describe a

smaller angle at the same angular frequency. Thus, the resonant variables oscillate at a frequency ω_o , which is larger than $\omega_n = 1$, i.e., $\omega_o > \omega_n$. Using (3.146), it is possible to verify that:

$$\omega_{nt} = \frac{\omega_n}{\sqrt{LC}} = \frac{1}{\sqrt{LC}}, \quad \omega_{ot} = \frac{\omega_o}{\sqrt{LC}} > \omega_{nt},$$

where ω_{nt} and ω_{ot} stand respectively, for the resonance frequency and the operation frequency, expressed with respect to real time t . Hence, it is concluded that the circuit must operate at frequencies that are greater than the resonance frequency to deliver output voltages smaller than the power supply voltage $v_0 < E$, i.e., $z_3 < 1$. This is the reason why, when the transistors are activated according to $u = \text{sign}(s)$, the circuit in Figs. 2.34a and 2.35 is designated as DC-to-DC high-frequency series resonant power electronic converter. Further information on DC-to-DC resonant power electronic converters can be found in [8], where converters of this class (series and parallel) are designed, constructed, controlled, and experimentally tested, and [7], where the method used here to analyze series resonant converters drawing arcs of a circle is studied. On the other hand, the use of the phase plane $z_2 - z_1$ to study the operation of resonant power converters was introduced for the first time in [9] and [10].

3.10 Summary

The systems to be controlled in engineering, and every component in a closed-loop control system, are described by ordinary, linear differential equations with constant coefficients. Hence, for a control engineer, the study of differential equations must be oriented toward the understanding of the properties of the differential equations that shape the solution waveform.

The solution of an ordinary, linear differential equation with constant coefficients is given as the addition of the natural and the forced responses. The natural response only depends on the roots of the differential equation characteristic polynomial and it is always present, even when the excitation function (or input) is zero. On the other hand, the forced response depends on the excitation function (or input), i.e., the forced response is zero if the excitation function is zero. The forced response can be rendered to be equal to the excitation function or input. Furthermore, if the differential equation is stable, i.e., if the natural response tends toward zero, the complete response converges to the forced response. Hence, the input of a differential equation can be chosen to be equal to the way in which it is desired that the differential equation solution behaves.

Thus, the rationale behind control system design is the following. Given a system or plant to be controlled, i.e., a differential equation, choose a controller, i.e., another differential equation, such that when feedback connects the two differential equations, a new equivalent closed-loop differential equation is obtained with the

following properties: 1) it is stable, 2) the forced response is equal to the input applied to the closed-loop system, i.e., the desired input of the closed-loop system, and 3) the natural response vanishes as fast and damped as desired.

The features listed in the previous paragraph are accomplished as follows:

1. Stability is determined exclusively by the roots of the differential equation characteristic polynomial or, equivalently, by the poles of the corresponding transfer function (see Sect. 3.4).
2. The steady-state response is achieved by rendering the forced response equal to the input applied to the closed-loop system. It is explained in Sect. 3.4 that this depends on the independent terms of polynomials at the numerator and denominator in the closed-loop transfer function for the case of constant inputs.
3. The transient response refers to the suitable shaping of the natural response waveform and is achieved by suitably assigning the closed-loop poles.

Controller design for arbitrary plants is studied in Chaps. 5, 6, 7 using the root locus method in addition to the frequency response and the state space approach. The main idea is to render the closed-loop system stable, so that the steady-state response is equal to the closed-loop system input and the transient response suitably shaped.

3.11 Review Questions

1. Why is a transfer function unstable if it has one pole with a zero real part that is repeated two or more times?
2. It has been said that the forced response has the same waveform as the input. Explain why this is not ensured if the transfer function has poles with a zero real part. Give an example of an input and a transfer function where this happens.
3. When an incandescent lamp is turned on it becomes hot. However, its temperature does not grow to infinity, but it stops at a certain value. Of the differential equations that have been studied in this chapter, which is the one that better describes the evolution in time of the incandescent lamp temperature? Why?
4. What is the relationship between a transfer function and a differential equation?
5. What are the conditions that determine the stability of a transfer function?
6. What are the necessary assumptions for a permanent magnet brushed DC motor to behave as an integrator or a double integrator?
7. How do you think the natural response waveform of a second-order differential equation whose characteristic polynomial has two real, repeated, and negative roots?
8. It has been explained that the time functions composing the natural response are determined by the transfer function poles, i.e., roots of the characteristic polynomial, but what is the effect of the transfer function zeros? What do you think determines the coefficients when using partial fraction expansion to apply the inverse Laplace transform? (see Sects. 3.1 and 3.5).

9. Indicate the zones of the s complex plane where the poles have to be located to achieve (i) fast but not oscillatory responses, (ii) fast and oscillatory responses, (iii) permanent oscillatory responses, and (iv) unstable responses.

3.12 Exercises

1. Consider the water level system in Example 3.2. The response in Fig. 3.43 is obtained when a constant water flow q_i is applied and the tank has a constant cross-section of $0.5[\text{m}^2]$. Find the values of both R and q_i .
2. Consider a proportional level control, i.e., when $q_i = k_1 k_p (h_d - h)$, with h_d the constant desired level. Using the results in the previous exercise to find the gain k_p , such that the steady-state error is less than or equal to 20% of the desired value and that the time constant is less than or equal to one third of the time constant in the previous exercise, i.e., the open-loop time constant. Find the maximal value of q_i that must be supplied. This is important for selecting the water pump to be employed. Suppose that the opening of the output valve is increased such that its hydraulic resistance decreases to 50%. Find the increment or the decrement, as a percentage, of the level steady-state error and the new time constant.

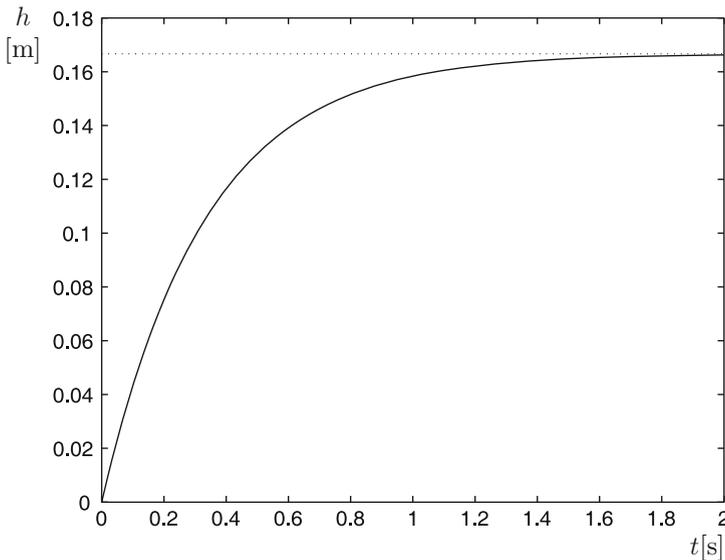


Fig. 3.43 Water level when a constant water flow is applied

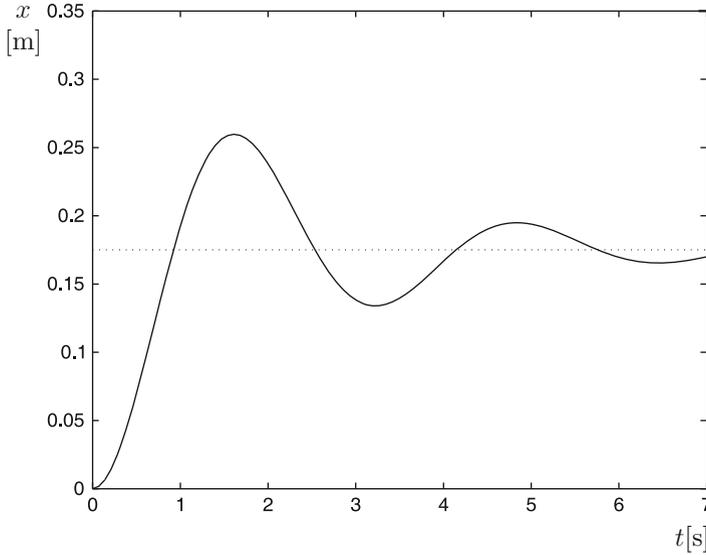


Fig. 3.44 Mass position in a mass-spring-damper system when a constant force is applied at the input. The dotted line stands for the final value of the mass position

3. Consider a proportional–integral level control. Show that oscillations in the water level appear if an adequate gain $k_i > 0$ is employed. Use your everyday experience to explain this phenomenon.
4. Given the following differential equation:

$$\ddot{y} + 127\dot{y} + 570y = 3990u, \quad y(0) = 0, \quad \dot{y}(0) = 0, \quad u = 1,$$

find values of ζ , ω_n and k in (3.53). With these values, write an expression for $y(t)$.

5. Consider the mass-spring-damper system described in Example 3.8. The response shown in Fig. 3.44 is obtained when the constant force $f=0.5[\text{N}]$ is applied. Find the values of b , K , and m .
6. Consider the following transfer functions:

$$Y_1(s) = G_1(s)U(s), \quad Y_2(s) = G_2(s)U(s),$$

$$G_1(s) = \frac{ab}{(s+a)(s+b)}, \quad G_2(s) = \frac{b}{s+b}, \quad a > 0, \quad b > 0,$$

where $U(s) = A/s$, with A a constant. Perform simulations where a takes values that increase from small to large (with respect to b) and compare graphically $y_1(t)$ and $y_2(t)$. Explain what is observed.

7. According to Sect. 3.1, it is known that from any initial condition the following differential equation $\dot{y} + 7y = 2$ has a solution given as $y(t) = Ee^{-7t} + D$ where E and D are some constants that are not going to be computed in this exercise. Proceeding similarly, find the solution for each one of the following differential equations.

$$\ddot{y} + 4\dot{y} + y = 2,$$

$$-\ddot{y} + \dot{y} + 10y = 3t,$$

$$\ddot{y} + 7\dot{y} = 2 \sin(t),$$

$$\ddot{y} + 3\dot{y} + 2y = 2e^{-t},$$

$$\ddot{y} + 2\dot{y} + 2y = 2 \cos(2t) - 4 \sin(2t),$$

$$\ddot{y} - 4y = 8t^2 - 4,$$

$$y^{(4)} + 11y^{(3)} + 28\ddot{y} = 10.$$

8. For each one of the following differential equations, find the value of $\lim_{t \rightarrow \infty} y(t)$. Notice that the same result can be obtained, no matter what the initial conditions are.

$$y^{(3)} + \ddot{y} - 2\dot{y} + y = 8,$$

$$\ddot{y} + \dot{y} - 10y = \cos(5t),$$

$$\ddot{y} + a\dot{y} + by = kA, \quad a > 0, b > 0, a, b, A, k \text{ are constants.}$$

9. Which one of the following differential equations has a solution with a smaller overshoot and which one has a solution with a smaller rise time?

$$3\ddot{y} + 3\dot{y} + 6y = 2A,$$

$$\ddot{y} + 4\dot{y} + 10y = 9A,$$

A is a constant.

10. Compute the poles of a second-order system with a 25% overshoot and 0.2[s] rise time.
11. Consider the proportional position control with velocity feedback of the following DC motor: $\ddot{\theta} + 0.4\dot{\theta} = 30u$, where θ is the angular position, whereas u is the applied voltage. Find the controller gains to achieve a 0.01[s] rise time and a 10% overshoot.
12. Consider the solution of a second-order differential equation when a step input is applied and assuming zero initial conditions, which is presented in (3.57). Find the maxima and minima of this response and solving for the first maximum, show that overshoot can be computed as:

$$M_p(\%) = 100 \times e^{-\frac{\xi}{\sqrt{1-\xi^2}}\pi}.$$

Find the times when the natural response is zero and solving for the first of these times, show that:

$$t_r = \frac{1}{\omega_d} \left[\pi - \arctan \left(\frac{\sqrt{1 - \zeta^2}}{\zeta} \right) \right].$$

13. Consider the following expressions:

$$Y(s) = \frac{as + b}{s^2 + c} U(s), \quad Z(s) = \frac{d}{s^2 + c} U(s),$$

where $U(s) = A/s$ with A, a, b, c, d , are positive constants. Express $y(t)$ as a function of $z(t)$.

14. Consider the following transfer function:

$$G(s) = \frac{c}{s^2 + bs + c}.$$

What must be done with b or c to achieve a faster system response?, What must be done with b or c to achieve a less oscillatory response?

15. Consider the solution $y(t)$ of an arbitrary n -order linear differential equation with constant coefficients.

- If the input is bounded, What are the conditions to ensure that $y(t)$ is also bounded for all time? i.e., to ensure that $y(t)$ does not become infinity.
- What are the conditions for the forced response to be the only component of $y(t)$ that remains as time increases?
- Assume that the input is zero. What are the different behaviors for $y(t)$ as time increases and under what conditions do they appear?
- Assume that the input is a polynomial of time. Show that the solution is not a polynomial with a larger degree than the polynomial at the input if all the roots of the characteristic polynomial have a negative real part.
- Assume that the input is given as the addition of several sine and cosine functions of time with different frequencies. Using the results in Sect. 3.6 and the superposition principle, Can you explain the following fundamental result for linear differential equations? [6, pp. 389]: “if the input is any periodic function with period T and all the roots of the characteristic polynomial have negative real parts, then as $t \rightarrow \infty$ the solution $y(t)$ is also a periodic function with period T , although with a waveform that may be different from that of the input.” Recall the concept of the Fourier series.

16. Verify the following equations:

$$i) \frac{5}{s^2 + s - 2} = \frac{-5/3}{s + 2} + \frac{5/3}{s - 1},$$

$$\begin{aligned}
 \text{ii)} \quad & \frac{2}{s(s+2)(s-1)} = \frac{1/3}{s+2} + \frac{2/3}{s-1} + \frac{-1}{s}, \\
 \text{iii)} \quad & \frac{1}{s^3 - s^2 - s + 1} = \frac{1/2}{(s-1)^2} + \frac{-1/4}{s-1} + \frac{1/4}{s+1}, \\
 \text{iv)} \quad & \frac{1}{s^4 - s^3 - s^2 + s} = \frac{1}{s} + \frac{1/2}{(s-1)^2} + \frac{-3/4}{s-1} + \frac{-1/4}{s+1}, \\
 \text{v)} \quad & \frac{2}{s^3 + 2s} = \frac{1}{s} + \frac{-s}{s^2 + 2}, \\
 \text{vi)} \quad & \frac{3}{s^4 - 3s^3} = \frac{-1}{s^3} + \frac{-1/3}{s^2} + \frac{-1/9}{s} + \frac{1/9}{s-3}, \\
 \text{vii)} \quad & \frac{8}{s(s+4)(s-4)} = \frac{-1/2}{s} + \frac{1/4}{s+4} + \frac{1/4}{s-4}, \\
 \text{viii)} \quad & \frac{8}{s^2 - 16} = \frac{1}{s-4} + \frac{-1}{s+4}, \\
 \text{ix)} \quad & \frac{4}{s^4 + s^3 + 2s^2} = \frac{2}{s^2} + \frac{-1}{s} + \frac{s-1}{s^2 + s + 2}, \\
 \text{x)} \quad & \frac{4}{s^5 + 3s^4 + 4s^3 + 4s^2 + 3s} = \frac{2}{(s+1)^3} + \frac{2}{(s+1)^2} + \frac{1}{s+1} + \frac{-s-1}{s^2+1}, \\
 \text{xi)} \quad & \frac{3s^2 + 3s - 2}{s^3 + 2s^2} = \frac{-1}{s^2} + \frac{2}{s} + \frac{1}{s+2}, \\
 \text{xii)} \quad & \frac{s+2}{s^2(s+1)^2} = \frac{2}{s^2} + \frac{-3}{s} + \frac{1}{(s+1)^2} + \frac{3}{s+1}.
 \end{aligned}$$

17. An induction oven is basically an inductor L with a heat-resistant material inside containing some metal to melt down. A capacitor C is series-connected to the inductor to improve the power factor. A high-frequency alternating voltage is applied at the circuit terminals that induces eddy currents in metal, producing heat and, hence, melting it down.

This device constitutes a series RLC circuit where the resistance R is represented by the equivalent electrical resistance of metal to melt down. The high-frequency alternating voltage applied to the circuit is given as a square wave described as $u = E \operatorname{sign}(i)$, where i is the electric current through the circuit, E is a positive constant and $\operatorname{sign}(i) = +1$ if $i \geq 0$ or $\operatorname{sign}(i) = -1$ if $i < 0$.

- Assume that the initial electric current is zero. Show that, provided that the damping coefficient is less than unity, i.e., $0 < \zeta < 1$, the electric current through the inductor is zero periodically with period π/ω_d , where $\omega_d = \omega_n \sqrt{1 - \zeta^2}$, $\omega_n = \frac{1}{\sqrt{LC}}$ y $\zeta = \frac{R}{2L\omega_n}$.

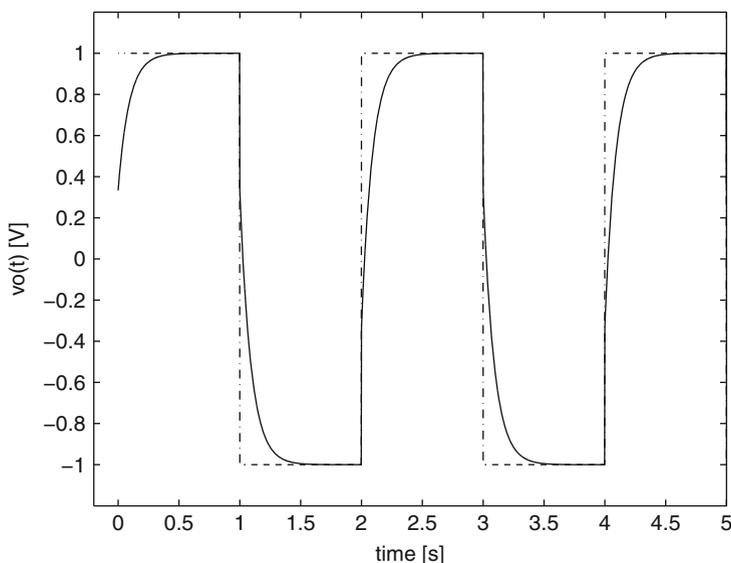
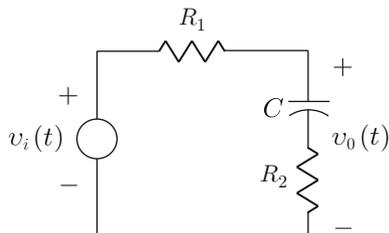
- According to $u = E \operatorname{sign}(i)$, the applied voltage, u , changes value when $i = 0$. Using this and $v_c(0)$ as the initial voltage at the capacitor, find an expression for the voltage at the capacitor that will be useful the next time that $i = 0$.
 - Propose any values for R , L , and C such that $0 < \zeta < 1$. Assuming that the voltage at the capacitor and the current through the inductor are not discontinuous when the value of u changes (why?), write a computer program to use iteratively the formula in the previous item to compute voltage at the capacitor after u has changed several times. Verify that the voltage at the capacitor, when $i = 0$, converges to a constant.
 - Perform a simulation of the series RLC circuit when $u = E \operatorname{sign}(i)$, to verify the result in the previous item.
18. According to Example 3.19, when the armature inductance is negligible, the mathematical model of a DC motor is given as:

$$\omega(s) = \frac{\frac{k_m}{JR}}{s + \left(\frac{B}{J} + \frac{k_m k_e}{JR_a}\right)} V(s),$$

where $\omega(s)$ and $V(s)$ are the Laplace transforms of the motor velocity and voltage applied at the motor terminals respectively.

- Assume that the applied voltage is zero and that the initial velocity is not zero. Notice that applying a zero voltage at motor terminals is equivalent to putting a short circuit at the armature motor terminals and that a non-zero-induced voltage is present if the velocity is not zero. Depict the natural response under these conditions. What can be done to force the natural response to vanish faster? What happens with the electric current? Can you explain why the natural response vanishes faster if the armature resistance is closer to zero? Do you know the meaning of the term “dynamic braking”? Why does the time constant depend on motor inertia? What does this mean from the point of view of Newton’s Laws of Mechanics?
 - Assume that a voltage different from zero is applied when the initial velocity is zero. Depict the system response. Why does the final velocity not depend on the motor inertia J ? Give an interpretation from the point of view of physics.
19. Consider the electric circuit in Fig. 3.45. This circuit is called a phase-lag network. The corresponding mathematical model is given in (2.146) and is rewritten here for ease of reference:

$$V_0(s) = \frac{R_2}{R_1 + R_2} \frac{s + \frac{1}{R_2 C}}{s + \frac{1}{(R_1 + R_2)C}} V_i(s) + \frac{R_1}{R_1 + R_2} \frac{1}{s + \frac{1}{(R_1 + R_2)C}} v_c(0),$$

Fig. 3.45 Phase-lag network**Fig. 3.46** Time response of the circuit in Fig. 3.45

where $v_c(0)$ is the initial voltage at the capacitor. Show that the solution is given as:

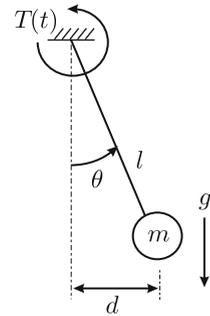
$$v_0(t) = A + \frac{R_1}{R_1 + R_2}(v_c(0) - A)e^{-at}, \quad a = \frac{1}{(R_1 + R_2)C},$$

and use this result to explain the response in Fig. 3.46 when $v_i(t)$ is the square wave represented by the dash-dot line, $v_c(0) = 0$, $R_1 = 2[\Omega]$, $R_2 = 1[\Omega]$ and $C = 1/33[\text{F}]$.

20. Repeat the previous exercise for all circuits in Fig. 2.39. Use MATLAB/Simulink to draw the time response and verify your results.
21. Consider the simple pendulum in Fig. 3.47 when the external torque $T(t) = 0$ and friction is negligible. The mathematical model is given as:

$$ml^2\ddot{\theta} + mgl\theta = 0,$$

Fig. 3.47 Simple pendulum



which is valid only when the pendulum remains close to the downward vertical configuration, i.e., θ is “small,” $\theta \approx 0$ (see Example 7.5, Chap. 7). Find the required pendulum length l to produce oscillations with a period of 1 s.

References

1. K. Ogata, *Modern control engineering*, 4th edition, Prentice-Hall, Upper Saddle River, 2002.
2. E. Kreyszig, *Advanced engineering mathematics*, Vol. 1, 3rd. edition, John Wiley and Sons, New Jersey, 1972.
3. C. Ray Wylie, *Advanced engineering mathematics*, McGraw-Hill, 1975.
4. M. R. Spiegel, *Mathematical handbook of formulas and tables*, McGraw-Hill, Schaum series, New York, 1974.
5. W. H. Hayt and J. E. Kemmerly, *Engineering circuit analysis*, McGraw-Hill, 1971.
6. C.-T. Chen, *Linear system theory and design*, Holt, Rinehart, and Winston, New York, 1984.
7. V. Hernández Guzmán, A new stability analysis method for DC to DC series resonant converters, *Computación y Sistemas*, vol. 11, no. 1, pp. 14–25, 2007.
8. R. Silva-Ortigoza, Control of resonant power converters based on differential flatness: design and construction, *MSc. Thesis*, CINVESTAV-IPN (Zacatenco), México City, February, 2002.
9. R. Oruganti and F.C. Lee, Resonant power processors, Part I—State plane analysis, *IEEE Trans. Industry Applications*, vol. AI-21, no. 6, pp. 1453–1460, 1985.
10. R. Oruganti and F.C. Lee, Resonant power processors, Part II—Methods of control, *IEEE Trans. Industry Applications*, vol. AI-21, no. 6, pp. 1461–1471, 1985.
11. M.E. Van Valkenburg, *Network analysis*, Prentice-Hall, Upper-Saddle River, 1974.
12. J. Stewart, *Calculus concepts and contexts*, Brooks Cole Publishing and IPT Company, 1998.