Chapter 2
Tools—Mainly Mathematics

For building models of physical systems we need mathematical tools. In order to model physical phenomena we have to look at lines, surfaces, and vector fields in three-dimensional space. The notion of a signal allows us to concentrate on the information contained in some time-varying physical quantity without thinking about the physical representation. The notion of a system likewise describes the information processing performed by some apparatus without resorting to the physical properties of the apparatus. We assume the reader is familiar with the material. Therefore, we quickly list the definitions and properties we need in the sequel. For a pedagogical treatment of the material see Lee and Varaiya (2011), Kreyszig (2010), Franklin et al. (2010).

For learning about inanimate nature we have to probe her and measure her response. In order to do so we must familiarize ourselves with instrumentation for generating, capturing, and analyzing real-world signals.

2.1 Complex Numbers

Let i denotes the imaginary unit, $\mathrm{i}^2 = -1$. A complex number $c \in \mathbb{C}$ is the sum $c = a + \mathrm{i}b$ for $a, b \in \mathbb{R}$. We denote the real part a of c with $\Re(c)$ and the imaginary part b with $\Im(c)$. The complex number $c^* = a - \mathrm{i}b$ denotes the complex conjugate of c. Euler's relation[1] (2.1) allows us to represent points on the unit-circle in a simple way,

$$\mathrm{e}^{\mathrm{i}\phi} = \cos\phi + \mathrm{i}\sin\phi, \tag{2.1}$$

where $\phi \in \mathbb{R}$. What makes Euler's relation so remarkable is that it links in a very useful way a purely algebraic concept—complex numbers—to geometry. Identifying the complex number $c = a + \mathrm{i}b = |c|\mathrm{e}^{\mathrm{i}\angle c}$ with the point $\vec{c} = (a, b)$ in the plane $|c|$ is

[1] Richard P. Feynman, Nobel laureate in Physics 1965, called Euler's relation the most remarkable formula in mathematics. We better remember this formula.

P. Hintenaus, *Engineering Embedded Systems*, DOI 10.1007/978-3-319-10680-9_2

the distance between the origin and the point $\vec{c}$. The angle $\angle c$ is the angle between the x-axis and the line through the origin and $\vec{c}$. Multiplying c and $\mathrm{e}^{\mathrm{i}\phi}$ amounts to rotating the point $\vec{c}$ counterclockwise by the angle ϕ around the origin, $c\mathrm{e}^{\mathrm{i}\phi} = |c|\mathrm{e}^{\mathrm{i}\angle c}\mathrm{e}^{\mathrm{i}\phi} = |c|\mathrm{e}^{\mathrm{i}(\angle c+\phi)} = (a+\mathrm{i}b)(\cos\phi + \mathrm{i}\sin\phi) = (a\cos\phi - b\sin\phi) + \mathrm{i}(a\sin\phi + b\cos\phi)$.

Whenever you encounter some identity involving complex numbers, which is not immediately obvious, you can most probably derive it from Euler's relation.

2.2 Line and Surface Integrals

We consider a curve C in space with the endpoints $\vec{a}$ and $\vec{b}$. Let C have a parametric representation $\vec{r}\colon \{t \in \mathbb{R}\colon a \le t \le b\} \to \mathbb{R}^3$,

$$t \mapsto \vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

for $a, b \in \mathbb{R}$, such that $\vec{r}(a) = \vec{a}$ and $\vec{r}(b) = \vec{b}$. To each value $a \le t_0 \le b$ corresponds a point of C whose position vector is $\vec{r}(t_0)$. The parametric representation imparts a direction onto C, from the start point $\vec{r}(a)$ to the end point $\vec{r}(b)$. If the curve C has a parametric representation $\vec{r}$ such that $\vec{r}(t_0)$ is continuous and has a continuous derivative $\vec{r}\,'(t_0) = \frac{\mathrm{d}\vec{r}}{\mathrm{d}t}(t_0)$, which is not identical to the zero-vector for $a \le t_0 \le b$, then the curve C has a unique tangent direction at any point, which varies continuously along C; then we declare the curve to be smooth.

Vector fields are functions that assign a vector to every point in space. Let $\vec{F}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be a continuous vector field. The line integral of $\vec{F}$ along the smooth curve C having a representation $\vec{r}$ with the properties above is

$$\int\limits_C \vec{F} \cdot \mathrm{d}\vec{l} = \int\limits_a^b \vec{F}(\vec{r}(t)) \cdot \frac{\mathrm{d}\vec{r}(t)}{\mathrm{d}t}\,\mathrm{d}t.$$

A piecewise smooth curve C consists of a finite number of smooth curves $C_1, \ldots, C_n$ such that the end point of C_i and the start point of C_{i+1} coincide for $1 \le i < n$. The line integral along C is the sum of the line integrals along the curves C_i. A curve is closed if its start point $\vec{a}$ and its endpoint $\vec{b}$ coincide. The line integral along a closed curve is denoted by $\oint_C \vec{F} \cdot \mathrm{d}\vec{l}$.

Let us consider now a surface S in three-dimensional space. Let S have a parametric representation $\vec{r}\colon R \to \mathbb{R}^3$,

$$(u, v) \mapsto \vec{r}(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix},$$

associating with every point (u_0, v_0) in some connected point set R of the uv-plane a point of S whose position vector is $\vec{r}(u_0, v_0)$. The point set R is connected if and only if any two points $\vec{p}_1, \vec{p}_2 \in R$ can be connected by finitely many line segments with each line segment contained entirely in R. The boundary B of the surface S is the set of all points p such that every neighborhood of p contains points in S as well as points not in S.

If the parametric representation $\vec{r}(u, v)$ is continuous and has continuous partial derivatives $\frac{\partial \vec{r}}{\partial u}(u_0, v_0)$ and $\frac{\partial \vec{r}}{\partial v}(u_0, v_0)$ with

$$\vec{N}(u_0, v_0) = \left(\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right)(u_0, v_0) \neq \vec{0}$$

for all points $(u_0, v_0) \in R$ then the surface S has at any point $\vec{r}(u_0, v_0)$ a unique tangent plane and a unique normal whose direction is given by $\vec{N}(u_0, v_0)$, both varying continuously across S; the surface is smooth. We denote the normal vector with unit length as $\vec{n} = \vec{N}/|\vec{N}|$.

A smooth surface S is orientable if and only if the positive normal direction at an arbitrary point $\vec{p}$ of S can be continued in a unique and continuous way to the whole surface. Let $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ be a continuous vector field. Let $\vec{r}$ be a parametric representation with the properties stated above of the smooth orientable surface S. The integral of $\vec{F}$ through the surface S is

$$\iint\limits_S \vec{F} \cdot \vec{n}\, \mathrm{d}A = \iint\limits_R \vec{F}(\vec{r}(u, v)) \cdot \left(\frac{\partial \vec{r}(u, v)}{\partial u} \times \frac{\partial \vec{r}(u, v)}{\partial v}\right) \mathrm{d}u\, \mathrm{d}v.$$

The Möbius strip is an example for a smooth surface that cannot be oriented, see e.g. (Kreyszig 2010).

A piecewise smooth surface S consists of finitely many smooth surfaces $S_1, \ldots, S_n$ and their boundaries. A piecewise smooth surface is orientable if and only if each surface S_i is orientable and for each pair of surfaces, which have part of their boundaries in common, S_i and S_j with $i \neq j$, the directions at the common part of the boundaries run against each other. The direction of the boundary of the surface S_i is set so that the normal and the boundary are right handed.

2.3 Discrete-Time Signals and Systems

A discrete-time signal s is defined at equally spaced times $t_n = n\sigma$, $n \in \mathbb{Z}$, where $\sigma \in \mathbb{R}, \sigma > 0$ is the time-step. It maps the t_n to the elements of some set $\mathbb{A}$. Therefore, the discrete-time signal s can be identified with the function $s : \mathbb{Z} \to \mathbb{A}, n \mapsto s_n$. We use subscript notation for discrete time signals throughout this book. Sometimes, it is convenient to restrict the domain of a discrete-time signal to the nonnegative integers. The signal s is periodic with period $p \in \mathbb{N}$ if and only if $s_n = s_{n+p}$ for all n in the domain of s. Most often the image of the signal s comprises either real or complex numbers, but more complicated situations arise. A video, for example, maps time to pictures. A picture in turn may be represented by a tuple of three two-dimensional arrays of real numbers, one array each for the colors red, green, and blue. In this book we restrict ourselves to real-valued, $\mathbb{A} = \mathbb{R}$, or complex-valued, $\mathbb{A} = \mathbb{C}$, signals.

The signal $\delta \colon \mathbb{Z} \to \mathbb{R}$,

$$n \mapsto \delta_n = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n \neq 0, \end{cases}$$

is called the discrete impulse. For an arbitrary signal $x \colon \mathbb{Z} \to \mathbb{C}$ the sifting property of the discrete impulse,

$$\sum_{n'=-\infty}^{\infty} x_{n'} \delta_{n-n'} = x_n, \tag{2.2}$$

holds for all $n \in \mathbb{Z}$. This property is easy to argue, but its frequent occurrence earns it a name.

The unit-step signal, $u \colon \mathbb{Z} \to \mathbb{R}$ is the sum of the discrete impulse,

$$n \mapsto u_n = \sum_{n'=-\infty}^{n} \delta_{n'} = \begin{cases} 0 & \text{for } n < 0 \\ 1 & \text{for } n \geq 0. \end{cases}$$

A discrete-time system[2] S uniquely transforms the input signal $x \colon \mathbb{Z} \to \mathbb{R}$, $n \mapsto x_n$, into the output signal $y \colon \mathbb{Z} \to \mathbb{R}$, $n \mapsto y_n$. We denote this fact by writing

$$y_n = (S(x))_n \,.$$

While we have written the definition for real-valued inputs and outputs a definition with the obvious modifications will do for complex-values input and output signals.

Discrete-time systems model digital feedback-controllers, for example, quite well. Such a controller measures the response of the process it is supposed to control at equally spaced times. From each measurement it computes a control output, which acts on the process via an actuator shortly after the controller has taken the measurement.

2.3.1 Linear Time-Invariant Systems

The discrete-time system S is linear if and only if the system allows superposition, that is

---

[2] We restrict our discussion to single-input single-output systems.

$$(S(au + bv))_n = a\,(S(u))_n + b\,(S(v))_n$$

for all signals $u\colon \mathbb{Z} \to \mathbb{R}$ and $v\colon \mathbb{Z} \to \mathbb{R}$ and for all $a \in \mathbb{R}$ and $b \in \mathbb{R}$. The system S is time-invariant if and only if

$$y_{n-k} = (S(\tilde{x}))_n$$

for all signals $x\colon \mathbb{Z} \to \mathbb{R}$ and $y\colon \mathbb{Z} \to \mathbb{R}$ and for all $k \in \mathbb{Z}$ where $y_n = (S(x))_n$ and the signal $\tilde{x}$, $\tilde{x}_m = x_{m-k}$ is the time-shifted signal x. Obvious modifications cover the complex-valued case.

2.3.2 Impulse Response and Convolution

Let $u\colon \mathbb{Z} \to \mathbb{C}$ and $v\colon \mathbb{Z} \to \mathbb{C}$ be complex-valued discrete-time signals. The signal $u * v\colon \mathbb{Z} \to \mathbb{C}$,

$$n \mapsto (u * v)_n = \sum_{n'=-\infty}^{\infty} u_{n'} v_{n-n'},$$

is called the convolution of u with v. The convolution is commutative, $u * v = v * u$. The shifting property (2.2) becomes $x = x * \delta$.

The impulse response of the linear time-invariant discrete-time system S is $h\colon \mathbb{Z} \to \mathbb{R}$, $n \mapsto h_n = (S(\delta))_n$. The impulse response describes the system S fully; the response $y\colon \mathbb{Z} \to \mathbb{R}$ of S to an arbitrary input $x\colon \mathbb{Z} \to \mathbb{R}$ is the convolution of h with x,

$$n \mapsto y_n = (S(x))_n = (h * x)_n. \tag{2.3}$$

The obvious modifications cover complex-valued inputs and outputs. When the impulse response $h_{n'}$ is nonzero for some $n' < 0$, computing y_n in (2.3) requires knowledge of $x_{n'}$ for $n' > n$. Having to know the future in order to produce a valid answer in the present is beyond a humble engineer's abilities. We call the system S causal if and only if $h_{n'} = 0$ for all $n' < 0$. Noncausal systems are useful only for processing retrospectively signals that have been recorded in advance.

2.3.3 Circular Convolution

For periodic signals the infinite sum in the definition of the convolution is not only unnecessary, but it also introduces problems with convergence. The circular convolution circumvents these problems by summing over a single period only. Let $x\colon \mathbb{Z} \to \mathbb{C}$ and $y\colon \mathbb{Z} \to \mathbb{C}$ be complex-valued periodic discrete-time signals with period p. The signal

$$(x \circledast y)_n = \sum_{n'=0}^{p-1} x_{n'} y_{n-n'}$$

for $n \in \mathbb{Z}$ is called the circular convolution of x and y.

2.4 Continuous-Time Signals and Systems

A continuous-time signal s maps real numbers to the elements of some set $\mathbb{A}$. It usually is a function. The domain of s most often is understood as time but other interpretations are possible. In Chap. 11 we will encounter signals whose domain is one-dimensional space instead of time. The signal $s\colon \mathbb{R} \to \mathbb{A}$, $t \mapsto s(t)$ is periodic with period $p \in \mathbb{R}$, $p > 0$ if and only if $s(t + p) = s(t)$ for all $t \in \mathbb{R}$. We restrict ourselves again to real-valued or complex-valued signals. Although some continuous-time signals are not functions we will use function-notation throughout this book.

A signal, which is not a function, is the Dirac delta, $\delta(t)$. For all $t \in \mathbb{R}$, $t \neq 0$ the Dirac delta is zero, $\delta(t) = 0$ and for all real $\epsilon > 0$,

$$\int_{-\epsilon}^{\epsilon} \delta(t)\,\mathrm{d}t = 1.$$

For getting an intuitive understanding[3] let us consider a strike to some object. We can describe such a strike by a function $P\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto b(t)$. Let us assume the strike begins at time $-d/2$ and ends at time $d/2$ for some duration $d \in \mathbb{R}$, $d > 0$. Before the impact at time $-d/2$ the strike does not transfer energy to the object, $b(t) = 0$ for $t < -d/2$. After the impact, the power $P(t)$ delivered to the object at time t will rise sharply. We assume that the strike delivers maximum power at time 0. Between time 0 and time $d/2$ the power drops off to zero again. The total energy transferred by the strike is $\int_{-\epsilon}^{\epsilon} P(t)\mathrm{d}t$ for $\epsilon > d/2$. When we consider a sequence of strikes all delivering the energy of 1 J with shorter and shorter duration the Dirac delta will be the limit of this sequence for $d \to 0$.

The Dirac delta has the sifting property; for a function $x\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto x(t)$ being continuous at $t \in \mathbb{R}$

$$x(t) = \int_{-\infty}^{\infty} x(\tau)\delta(t - \tau)\,\mathrm{d}\tau \tag{2.4}$$

holds.

[3] A precise mathematical treatment involves the theory of distributions and measure theory. This is well beyond the scope of this book.

The unit-step signal, $u\colon \mathbb{R} \to \mathbb{R}$ is the integral of the Dirac delta,

$$t \mapsto u(t) = \int_{-\infty}^{t} \delta(\tau)\mathrm{d}\tau = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t \geq 0. \end{cases}$$

The instantaneous power $P(s(t))$ at time t of the complex valued signal $s\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto s(t)$ is $P(s(t)) = s(t)s(t)^*$. The average power $\mathcal{P}(s)$ of s is

$$\mathcal{P}(s) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t)s(t)^*\mathrm{d}t. \tag{2.5}$$

A continuous-time system[4] S uniquely transforms the input signal $x\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto x(t)$ into an output signal $y\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto y(t)$. We denote this fact by writing

$$y(t) = (S(x))\,(t).$$

The definition covers real-valued inputs and real-valued outputs. For complex-valued inputs and outputs obvious modifications apply.

Continuous-time systems model electrical networks quite well, see Chap. 3. Such a network with one input and one output reacts to a voltage at its input with a voltage at its output.

2.4.1 Linear Time-Invariant Systems

The continuous-time system S is linear if and only if the system allows superposition, that is

$$(S(au + bv))\,(t) = a\,(S(u))\,(t) + b\,(S(v))\,(t)$$

for all signals $u\colon \mathbb{R} \to \mathbb{R}$ and $v\colon \mathbb{R} \to \mathbb{R}$ and for all $a \in \mathbb{R}$ and $b \in \mathbb{R}$. The system S is time-invariant if and only if

$$y(t - t') = (S(\tilde{x}))\,(t)$$

for all signals $x\colon \mathbb{R} \to \mathbb{R}$ and $y\colon \mathbb{R} \to \mathbb{R}$ and for all $t' \in \mathbb{R}$ where $y(t) = (S(x))\,(t)$ and $\tilde{x}(t) = x(t - t')$. Again the obvious modifications will cover the complex-valued case.

[4] We again restrict our discussion to single-input single-output systems.

2.4.2 Impulse Response and Convolution

Let $x: \mathbb{R} \to \mathbb{C}$ and $y: \mathbb{R} \to \mathbb{C}$ be complex-valued signals. The signal

$$(x * y)(t) = \int_{-\infty}^{\infty} x(\tau)y(t-\tau)\,\mathrm{d}\tau$$

is called the convolution of x with y. Equation (2.4) becomes $x = x * \delta$.

The impulse response of the linear time-invariant continuous-time system S is $h: \mathbb{R} \to \mathbb{R}$, $t \mapsto h(t) = (S(\delta))\,(t)$. The impulse response describes the system S fully; the response $y: \mathbb{R} \to \mathbb{R}$ of S to an arbitrary input $x: \mathbb{R} \to \mathbb{R}$ is the convolution of h with x,

$$t \mapsto y(t) = (S(x))\,(t) = (h * x)(t). \tag{2.6}$$

When the impulse response $h(t')$ is nonzero for some $t' < 0$ computing $y(t)$ in (2.6) requires knowledge of $x(t')$ for $t' > t$. The future's not ours to see; therefore we call the system S causal if and only if $h(t') = 0$ for all $t' < 0$. Noncausal systems serve as theoretical tools only. The obvious modifications cover systems with complex valued inputs and outputs.

2.4.3 Circular Convolution

For periodic signals the improper integral in the definition of the convolution is not only unnecessary but it also introduces problems with convergence. The circular convolution circumvents these problems by integrating over a single period only. Let $x: \mathbb{R} \to \mathbb{C}$ and $y: \mathbb{R} \to \mathbb{C}$ be complex-valued periodic signals with period p. The signal

$$(x \circledast y)(t) = \frac{1}{p}\int_{0}^{p} x(\tau)y(t-\tau)\,\mathrm{d}\tau$$

is called the circular convolution of x and y.

2.5 The Four Fourier Transforms

Under very general conditions we can describe a signal s as the weighted sum of phase-shifted sinusoids using the Fourier transform appropriate for the type of signal at hand. This representation is called the spectrum of s. The distribution of the amplitude with respect to frequency is the amplitude spectrum of s; the distribution of the phase with respect to frequency is the phase spectrum of s. Distinguishing between discrete-time and continuous-time signals, and between periodic and aperiodic signals gives rise to four possibilities. Discrete-time signals have periodic Fourier transforms, continuous-time signals have aperiodic Fourier transforms. Periodic signals have discrete Fourier transforms and aperiodic signals have continuous Fourier transforms.

2.5.1 Periodic Continuous-Time Signals—The Fourier Series

Let $s\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto s(t)$ be a continuous complex-valued periodic signal with period $p \in \mathbb{R}$, that is $s(t + p) = s(t)$ for every real number t. The complex Fourier coefficients of s constitute the sequence $S\colon \mathbb{Z} \to \mathbb{C}$,

$$n \mapsto S_n = \frac{1}{p} \int_{-\frac{p}{2}}^{\frac{p}{2}} s(t)\mathrm{e}^{-\mathrm{i}n\omega_0 t}\, \mathrm{d}t. \tag{2.7}$$

where $\omega_0 = \frac{2\pi}{p}$. As the integrand is periodic with period p the bounds can be chosen freely, as long as the integration domain spans exactly one period. The signal s can be represented by its Fourier series

$$s(t) \sim \sum_{n=-\infty}^{\infty} S_n \mathrm{e}^{\mathrm{i}n\omega_0 t}, \tag{2.8}$$

regardless of convergence. If the signal s is piecewise continuous in the interval $-\frac{p}{2} \leq t \leq \frac{p}{2}$, i.e. s has only finitely many finite jumps in the interval, and s is of bounded variation in the interval, then equality holds in (2.8) for all t where s is continuous. At a point t_0, where s is discontinuous, the sum in (2.8) is the average of the left-hand and the right-hand limit of s at t_0. The conditions above are called the Dirichlet conditions. The variation $V_a^b(s)$ of s in the interval $[a, b] = \{x \in \mathbb{R}\colon a \leq x \leq b\}$ is defined as

$$V_a^b(s) = \sup_{\substack{m \in \mathbb{N} \\ t_1 < \cdots < t_m \in [a,b]}} \sum_{i=1}^{m-1} |s(t_{i+1}) - s(t_i)|.$$

The variation of $f(t) = \sin\frac{1}{t}$ for $t \neq 0$ and $f(0) = 0$ for example is not bounded in the interval $[0, 2\pi]$, the variation of $g(t) = t^2 \sin\frac{1}{t}$ in the same interval, however, is bounded (Table 2.1).

Table 2.1 Some continuous-time periodic signals and their Fourier series coefficients

$s(t)$	$S_n = (\mathcal{F}(s))_n$	Comment
$x(t)h(t)$	$(X * H)_n$	
$(x \circledast h)(t)$	$X_n H_n$	
$x(t)$ where x is real-valued	X_n where $X_n = X^*_{-n}$	
$x(t)$ where $x(t) = x(-t)^*$	X_n where X is real-valued	
$s(t) = \begin{cases} 1 & \text{if } \frac{-p}{4} \le t < \frac{p}{4} \\ -1 & \text{if } \frac{p}{4} \le t < \frac{3p}{4} \\ s(t-p) & \text{otherwise} \end{cases}$	$S_n = \begin{cases} 0 & \text{if } n \text{ even} \\ 2\frac{(-1)^{n/2+1}}{n\pi} & \text{if } n \text{ odd} \end{cases}$	Square wave
$s(t) = \begin{cases} t + \frac{p}{4} & \text{if } \frac{-p}{2} \le t < 0 \\ -t + \frac{p}{4} & \text{if } 0 \le t < \frac{p}{2} \\ s(t-p) & \text{otherwise} \end{cases}$	$S_n = \begin{cases} 0 & \text{if } n \text{ even} \\ \frac{p}{\pi^2 n^2} & \text{if } n \text{ odd} \end{cases}$	Triangle
$s(t) = \begin{cases} t & \text{if } \frac{-p}{2} \le t < \frac{p}{2} \\ s(t-p) & \text{otherwise} \end{cases}$	$\mathrm{i}p\frac{(-1)^n}{2n\pi}$	Sawtooth

The periodic signals with period p, $x\colon \mathbb{R} \to \mathbb{C}$ and $h\colon \mathbb{R} \to \mathbb{C}$, are arbitrary. Their Fourier coefficients constitute the sequences $X\colon \mathbb{Z} \to \mathbb{C}$ and $H\colon \mathbb{Z} \to \mathbb{C}$ respectively

For a periodic real-valued signal $s\colon \mathbb{R} \to \mathbb{R}$ the following notation often is more convenient. Using $\cos\phi = \frac{1}{2}(\mathrm{e}^{\mathrm{i}\phi} + \mathrm{e}^{-\mathrm{i}\phi})$ and $\sin\phi = \frac{\mathrm{i}}{2}(\mathrm{e}^{-\mathrm{i}\phi} - \mathrm{e}^{\mathrm{i}\phi})$ we can rewrite (2.8) as

$$s(t) \sim a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t), \tag{2.9}$$

where

$$a_0 = \frac{1}{p} \int_{-\frac{p}{2}}^{\frac{p}{2}} s(t)\,\mathrm{d}t,$$

$$a_n = \frac{2}{p} \int_{-\frac{p}{2}}^{\frac{p}{2}} s(t) \cos n\omega_0 t\,\mathrm{d}t,$$

$$b_n = \frac{2}{p} \int_{-\frac{p}{2}}^{\frac{p}{2}} s(t) \sin n\omega_0 t\,\mathrm{d}t.$$

The frequency $\frac{\omega_0}{2\pi}$ is called the fundamental frequency or the first harmonic, while the frequency $\frac{n\omega_0}{2\pi}$ is called the frequency of the $(n-1)$st overtone, or the nth harmonic. We can rewrite (2.9) further into a form containing amplitudes and phases,

$$s(t) \sim A_0 + \sum_{n=1}^{\infty} A_n \cos\left(n\omega_0 t - \phi_n\right),$$

where the offset A_0 is $A_0 = a_0$, the amplitude A_n of the nth harmonic is $A_n = \sqrt{a_n^2 + b_n^2}$ and the phase ϕ_n of the nth harmonic is $\phi_n = \arctan \frac{b_n}{a_n}$.

2.5.2 Periodic Discrete-Time Signals—the Discrete Fourier Transform

Let $s\colon \mathbb{Z} \to \mathbb{C}$, $n \mapsto s_n$ be a discrete-time complex-valued periodic signal with period p. The coefficients of the discrete Fourier transform $\mathcal{F}(s)$ of s constitute the periodic sequence $S\colon \mathbb{Z} \to \mathbb{C}$,

$$k \mapsto S_k = \sum_{n=0}^{p-1} s_n \mathrm{e}^{-\mathrm{i}n\omega_0 k}, \tag{2.10}$$

with period p, where $\omega_0 = \frac{2\pi}{p}$. The signal s can be represented in terms of the Fourier transform coefficients by

$$s_n = \frac{1}{p} \sum_{k=0}^{p-1} S_k \mathrm{e}^{\mathrm{i}k\omega_0 n}. \tag{2.11}$$

As the summands in both sums are periodic with period p, the bounds in both sums can be chosen freely, as long as the summations cover exactly one period. For a real-valued signal s we can rewrite this equation for odd periods p as

$$s_n = \frac{1}{p} S_0 + \frac{2}{p} \sum_{k=1}^{(p-1)/2} \Re(S_k) \cos k\omega_0 n - \Im(S_k) \sin k\omega_0 n, \tag{2.12}$$

using $\mathrm{e}^{\mathrm{i}\phi} = \cos\phi + \mathrm{i}\sin\phi$ and $S_{-k} = S_k^*$.

When measuring a real-valued signal s' it is often a perturbed version of a true signal s. The perturbations, called noise, can be added for example by the inevitable imperfections of the measurement apparatus used for capturing s'. Let $s'_n = s_n + x_n$ where the x_n are normal distributed independent random variables all with zero mean and variance σ^2. The real part $\Re(S'_k)$ and the imaginary part $\Im(S'_k)$ of each Fourier transform coefficient S'_k then are random variables with mean $\Re(S_k)$ and

Table 2.2 Some discrete-time periodic signals and their discrete Fourier transforms

s_n	$S_k = (\mathcal{F}(s))_k$
$x_n h_n$	$\frac{1}{p}\sum_{k'=0}^{p-1} X_{k'} H_{k-k'} = \frac{1}{p}(X \circledast H)_k$
$(x \circledast h)_n$	$X_k H_k$
x_n where x is real-valued	X_k where $X_k = X^*_{-k}$
x_n where $x_n = x^*_{-n}$	X_k where X is real-valued
$e^{i2\pi f n}$ where $f = m/p, m \in \mathbb{Z}, m \neq 0$	$S_k = \begin{cases} 1 & \text{if } k \in \{m + ip : i \in \mathbb{Z}\} \\ 0 & \text{otherwise} \end{cases}$

The discrete-time periodic signals with period p, $x : \mathbb{Z} \to \mathbb{C}$ and $h : \mathbb{Z} \to \mathbb{C}$, are arbitrary, $X : \mathbb{Z} \to \mathbb{C}$ and $H : \mathbb{Z} \to \mathbb{C}$ are their discrete Fourier transforms

$\Im(S_k)$ respectively, and variance $\frac{p}{2}\sigma^2$, (Schoukens and Renneboog 1986). When reconstructing the signal s'_n according to (2.12) from the Fourier transform coefficients S'_k the real and imaginary parts of the Fourier transform coefficients are scaled with $2/p$. These scaled values then have variance $\frac{2}{p}\sigma^2$. The ratio $\frac{p}{2}$ of the variance of the s'_n and the variance of the real and imaginary parts of the scaled S'_k is called the processing gain of the discrete Fourier transform (Table 2.2).

2.5.3 *Discrete-Time Signals—The Discrete-Time Fourier Transform*

Let $s : \mathbb{Z} \to \mathbb{C}$, $n \mapsto s_n$ be a discrete complex-valued signal. The discrete-time Fourier transform $\mathcal{F}(s)$ of s is the function $S : \mathbb{R} \to \mathbb{C}$,

$$\omega \mapsto S(\omega) = \sum_{n=-\infty}^{\infty} s_n e^{-i\omega n}. \tag{2.13}$$

If s is absolutely summable, i.e., $\sum_{n=-\infty}^{\infty} |s_n|$ converges to a finite value, then the discrete-time Fourier transform exists and is finite for all ω. The transform is continuous and periodic with period 2π. The discrete signal s can be recovered by the inverse transform $\mathcal{F}^{-1}(S)$,

$$s_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) e^{i\omega n}\, d\omega. \tag{2.14}$$

Table 2.3 Some discrete-time signals and their discrete-time Fourier transforms

s_n	$S(\omega) = (\mathcal{F}(s))(\omega)$
$x_n h_n$	$\frac{1}{2\pi}\int_0^{2\pi} X(\Omega)H(\omega-\Omega)\,\mathrm{d}\Omega = (X \circledast H)(\omega)$
$(x * h)_n$	$X(\omega)H(\omega)$
x_n where x is real-valued	$X(\omega)$ where $X(\omega) = X(-\omega)^*$
x_n where $x_n = x^*_{-n}$	$X(\omega)$ where X is real-valued
$\delta_{n-n'}$	$\mathrm{e}^{-\mathrm{i}\omega n'}$
$\mathrm{e}^{\mathrm{i}\omega_0 n}$	$2\pi\,\delta(\omega-\omega_0)$

The discrete-time signals $x\colon \mathbb{Z} \to \mathbb{C}$ and $h\colon \mathbb{Z} \to \mathbb{C}$ are arbitrary, $X\colon \mathbb{R} \to \mathbb{C}$ and $H\colon \mathbb{R} \to \mathbb{C}$ are their discrete-time Fourier transforms

As the integrand is periodic with period 2π, one can choose the bounds freely, as long as the integration spans one period.

A finite discrete signal $s\colon \{0,\dots,p-1\} \to \mathbb{C}$ can be extended to the signal $s'\colon \mathbb{Z} \to \mathbb{C}$ by setting $s'_n = s_n$ for $n \in \{0,\dots,p-1\}$ and $s'_n = 0$ for $n \notin \{0,\dots,p-1\}$, or to the signal $s''\colon \mathbb{Z} \to \mathbb{C}$ by extending it periodically, i.e. $s''_n = s_{n \bmod p}$. Then the discrete Fourier transform coefficients S''_k of s'' are samples of the discrete-time Fourier transform S' of s',

$$S''_k = S'\left(\frac{2\pi}{p}k\right). \tag{2.15}$$

2.5.3.1 Frequency Response of Discrete-Time Linear Time-Invariant Systems

When we subject the discrete-time linear time-invariant system S to the input signal $x\colon \mathbb{Z} \to \mathbb{C}$, $n \mapsto x_n = \mathrm{e}^{\mathrm{i}\omega n}$ for $\omega \in \mathbb{R}$, then according to (2.3) the answer is $h * x$ where h is the impulse response of S. Expanding the convolution we get

$$(x * h)_n = \sum_{n'=-\infty}^{\infty} h_{n'}\mathrm{e}^{\mathrm{i}\omega(n-n')} = \mathrm{e}^{\mathrm{i}\omega n}\sum_{n'=-\infty}^{\infty} h_{n'}\mathrm{e}^{-\mathrm{i}\omega n'}.$$

Provided the sum exists we can rewrite this as $(x * h)_n = H(\omega)\mathrm{e}^{\mathrm{i}\omega n}$ where H, the frequency response of the system S, is the discrete-time Fourier transform of h (Table 2.3).

The response to $x'\colon \mathbb{Z} \to \mathbb{R}$,

$$n \mapsto \cos\omega n = \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\omega n} + \mathrm{e}^{-\mathrm{i}\omega n}\right)$$

is $y' \colon \mathbb{Z} \to \mathbb{C}$,

$$n \mapsto y'_n = \frac{1}{2}\left(H(\omega)\mathrm{e}^{\mathrm{i}\omega n} + H(-\omega)\mathrm{e}^{-\mathrm{i}\omega n}\right).$$

If the impulse response h is real-valued then $H(-\omega) = H(\omega)^*$ and

$$y'_n = |H(\omega)| \cos\left(\omega n + \angle H(\omega)\right),$$

where $\angle H(\omega)$ is the phase response of the system S at the angular frequency ω. The response of a discrete-time linear time-invariant system with real-valued impulse response to a sinusoidal input is a sinusoidal output with the same frequency as the input but with shifted phase.

Let us consider the discrete-time system S described by the equation

$$y_n = x_n + \alpha y_{n-1} \tag{2.16}$$

for $0 < \alpha < 1$. Let $x_n = 0$ for $n < n'$ and some $n' \in \mathbb{N}$. Provided the outputs y_n are all zero for $n < n'$ the system is linear and time invariant. Its impulse response is $h \colon \mathbb{Z} \to \mathbb{R}$,

$$n \mapsto S(\delta)_n = h_n = \begin{cases} 0 & \text{for } n < 0 \\ \alpha^n & \text{for } n \geq 0. \end{cases}$$

Therefore, the response to an arbitrary signal $x \colon \mathbb{Z} \to \mathbb{R}$, $n \mapsto x_n$ is $y \colon \mathbb{Z} \to \mathbb{R}$

$$n \mapsto y_n = \sum_{k=0}^{\infty} x_{n-k}\alpha^k.$$

The frequency response is $H : \mathbb{C} \to \mathbb{C}$,

$$\omega \mapsto H(\omega) = \frac{\mathrm{e}^{\mathrm{i}\omega}}{\mathrm{e}^{\mathrm{i}\omega} - \alpha}.$$

For plotting the frequency response for one period, Fig. 2.1, we compute gain and phase of the frequency response, $H(\omega) = |H(\omega)|\mathrm{e}^{\mathrm{i}\angle H(\omega)}$ and plot the gain $|H(\omega)|$ and the phase $\angle H(\omega)$ separately. We plot the gain on a logarithmic scale. In electrical engineering the decibel, [dB], denotes a ratio of powers. A ratio of 100 corresponds to 10dB, a ratio of 2 to about 3dB. More precisely, the ratio of two powers P_1 and P_2 is

$$10 \log_{10} \frac{P_1}{P_2}\,\mathrm{dB}.$$

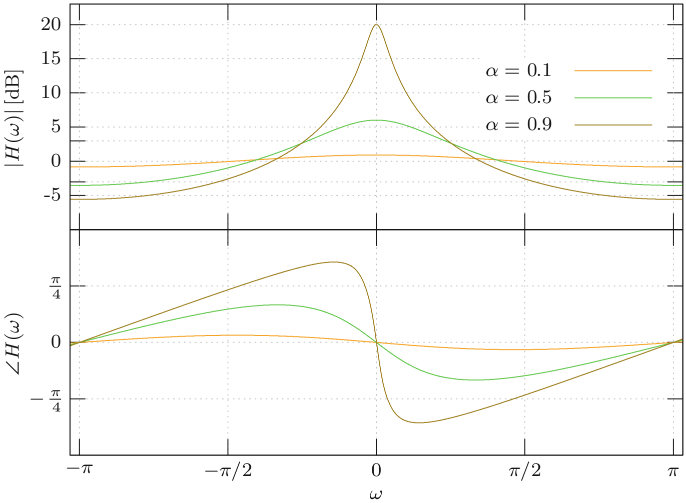


Fig. 2.1 Gain and phase of the discrete-time system described by Eq. (2.16)

Given two sinusoidal signals with voltage-amplitudes a_1 and a_2 we consider the ratio of the powers P_1 and P_2. The power P_1 is the power the first signal makes a resistor dissipate, while P_2 is the power the second signal makes the same resistor dissipate. A resistor is an electrical component with two terminals dissipating power P, which is proportional to V^2 when the voltage V is applied across it, see Sect. 3.3. Therefore, the power-ratio of the two sinusoidal signals is

$$10 \log_{10} \frac{P_1}{P_2} \text{ dB} = 10 \log_{10} \frac{a_1^2}{a_2^2} \text{ dB} = 20 \log_{10} \frac{a_1}{a_2} \text{ dB}.$$

2.5.4 Continuous-Time Signals—The Continuous-Time Fourier Transform

Let $s\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto s(t)$ be a continuous aperiodic complex-valued signal. The Fourier transform $\mathcal{F}(s)$ of s is the function $S\colon \mathbb{R} \to \mathbb{C}$,

$$\omega \mapsto S(\omega) = \lim_{T\to\infty} \int_{-T/2}^{T/2} s(t)\mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}t. \tag{2.17}$$

The transform is continuous and aperiodic. The signal s can be recovered by the reverse transform $\mathcal{F}^{-1}(S)$

$$s(t) \sim \frac{1}{2\pi} \lim_{\Omega\to\infty} \int_{-\Omega/2}^{\Omega/2} S(\omega)\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}\omega. \tag{2.18}$$

If the signal s is absolutely integrable, i. e., the integral $\int_{-\infty}^{\infty} |s(t)|\,\mathrm{d}t$ converges to a finite value, and s is piecewise continuous in every finite interval, and s is of bounded variation in every finite interval, then the equality in (2.18) holds for all t where s is continuous. At a point t_0, where s is discontinuous, the integral in (2.18) is the average of the left-hand and the right-hand limit of s at t_0.

If the signal s is periodic with period p having Fourier series coefficients S_n, then the continuous-time Fourier transform $S(\omega)$ of s is

$$S(\omega) = 2\pi \sum_{n=-\infty}^{\infty} S_n \delta\left(\omega - \frac{2\pi}{p} n\right). \tag{2.19}$$

Let $s' \colon \mathbb{Z} \to \mathbb{C}$ be a sampled version of the signal s, $n \mapsto s'_n = s(\tau n)$, where τ is the sample period. The discrete-time Fourier transform S' of s' is related to the Fourier transform S of s by

$$S'(\omega) = \frac{1}{\tau} \sum_{k=-\infty}^{\infty} S\left(\frac{\omega - 2\pi k}{\tau}\right). \tag{2.20}$$

If the signal s is band-limited such that $S(\omega) = 0$ for $\omega \le \frac{-\pi}{\tau}$ or $\omega \ge \frac{\pi}{\tau}$ the sum reduces to

$$S'(\omega) = \frac{1}{\tau} S\left(\frac{\omega}{\tau}\right).$$

2.5.4.1 Frequency Response of Continuous-Time Linear Time-Invariant Systems

When we subject the continuous-time linear time-invariant system S to the input signal $x \colon \mathbb{R} \to \mathbb{C}$, $t \mapsto x(t) = \mathrm{e}^{\mathrm{i}\omega t}$ for $\omega \in \mathbb{R}$, then according to (2.6) the answer is $h * x$ where h is the impulse response of S. Expanding the convolution we get $(x * h)(t) = \int_{-\infty}^{\infty} h(\tau)\mathrm{e}^{\mathrm{i}\omega(t-\tau)}\mathrm{d}\tau = \mathrm{e}^{\mathrm{i}\omega t} \int_{-\infty}^{\infty} h(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau}\mathrm{d}\tau$. Provided the integrals exist we can rewrite this as $(x * h)_n = H(\omega)\mathrm{e}^{\mathrm{i}\omega t}$ where H, the frequency response of the system S, is the continuous-time Fourier transform of h (Table 2.4).

The response to $x' \colon \mathbb{R} \to \mathbb{R}$, $t \mapsto \cos\omega t = \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\omega t} + \mathrm{e}^{-\mathrm{i}\omega t}\right)$ is $y' \colon \mathbb{R} \to \mathbb{C}$, $t \mapsto y'(t) = \frac{1}{2}\left(H(\omega)\mathrm{e}^{\mathrm{i}\omega t} + H(-\omega)\mathrm{e}^{-\mathrm{i}\omega t}\right)$. If the impulse response h is real-valued then $H(-\omega) = H(\omega)^*$ and

Table 2.4 Some continuous-time signals and their continuous-time Fourier transforms

$s(t)$	$S(\omega) = (\mathcal{F}(s))(\omega)$	Comment
$x(t)h(t)$	$\frac{1}{2\pi}\int_{-\infty}^{\infty} X(\Omega)H(\omega-\Omega)\,\mathrm{d}\Omega$ $= \frac{1}{2\pi}(X * H)(\omega)$	
$(x * h)(t)$	$X(\omega)H(\omega)$	
$\frac{\mathrm{d}x(t)}{\mathrm{d}t}$	$\mathrm{i}\omega X(\omega)$	
$x(t)$ where x is real-valued	$X(\omega)$ where $X(\omega) = X(-\omega)^*$	
$x(t)$ where $x(t) = x(-t)^*$	$X(\omega)$ where X is real-valued	
$\delta(t-T)$	$\mathrm{e}^{-\mathrm{i}\omega T}$	
$\mathrm{e}^{\mathrm{i}\Omega t}$	$2\pi\delta(\omega-\Omega)$	
$x(t-T)$	$\mathrm{e}^{-\mathrm{i}\omega T}X(\omega)$	
$\mathrm{e}^{\mathrm{i}\Omega t}x(t)$	$X(\omega-\Omega)$	
$s(t) = \begin{cases} 1 & \text{if } -p \le t \le p \\ 0 & \text{otherwise} \end{cases}$	$\frac{2\sin\omega p}{\omega}$	Boxcar
$s(t) = \begin{cases} t+p & \text{if } -p \le t < 0 \\ -t+p & \text{if } 0 \le t < p \\ 0 & \text{otherwise} \end{cases}$	$\frac{2-2\cos\omega p}{\omega^2}$	Triangle

The signals $x\colon \mathbb{R} \to \mathbb{C}$ and $h\colon \mathbb{R} \to \mathbb{C}$ are arbitrary, $X\colon \mathbb{R} \to \mathbb{C}$ and $H\colon \mathbb{R} \to \mathbb{C}$ are their Fourier transforms. The numbers $T \in \mathbb{R}$ and $\Omega \in \mathbb{R}$ are arbitrary, the number $p \in \mathbb{R}$, $p > 0$ is arbitrary

$$y'(t) = |H(\omega)|\cos(\omega t + \angle H(\omega)),$$

where $\angle H(\omega)$ is the phase response of the system S at the angular frequency ω. The response of a continuous-time linear time-invariant system with real-valued impulse response to a sinusoidal input is a sinusoidal output with the same frequency as the input but with shifted phase.

2.6 Noise

Any common signal processing apparatus introduces errors to the signals they process. A portion of these errors may be attributable to inadequacies of the apparatus; the aleatory behavior of nature at the microscopic level, however, introduces the rest. We call these errors noise. Any practical signal s will be degraded by noise. Most often the observed signal s will be the sum of the noiseless signal s' and some added noise n, $s(t) = s'(t) + n(t)$. The signal to noise ratio $\mathrm{SNR}(s)$ then is the ratio of the average power of s' to the average power of n,

$$\mathrm{SNR}(s) = \frac{\mathcal{P}(s')}{\mathcal{P}(n)}.$$

The signal to noise ratio is usually stated in decibel.

The spectral power density describes the frequency content of noise.[5] For each frequency f the spectral power density of a signal is the power the signal contains in a small frequency band around f. Band-limited white noise has a power spectral density being constant up to some band limit above which it drops to zero. Many electronic devices introduce noise with a spectral power density proportional to $1/f$ into the signals they process. These devices negatively affect low-frequency signals most, forcing us to make precision measurements at frequencies well away from 0Hz.

2.7 The Z-Transform

Let $x\colon \mathbb{Z} \to \mathbb{C}$, $n \mapsto x_n$ be a discrete-time signal. The Z-transform of x, $\hat{X}\colon \mathrm{roc}(x) \to \mathbb{C}$ is defined as

$$z \mapsto \hat{X}(z) = (\mathcal{Z}(x))\,(z) = \sum_{n=-\infty}^{\infty} z^{-n} x_n. \tag{2.21}$$

The set $\mathrm{roc}(x) \subseteq \mathbb{C}$ is the region of convergence of the Z-transform of x. When we substitute $z = r\mathrm{e}^{\mathrm{i}\omega}$ in (2.21) we recognize that the Z-transform is actually the discrete-time Fourier transform of the signal $x'\colon \mathbb{Z} \to \mathbb{C}$, $n \mapsto x'_n = r^{-n} x_n$. The sum in (2.21) converges exactly for those $r \in \mathbb{R}$ for which the sum $\sum_{n=-\infty}^{\infty} \left|r^{-n} x_n\right|$ converges to a finite value.

If x is right-sided, that is $x_n = 0$ for all $n < N$ and some $N \in \mathbb{Z}$, the region of convergence is the whole complex plane with the exception of a disc around the origin, $\mathrm{roc}(x) = \left\{z \in \mathbb{C}\colon |z| > r'\right\}$ for some real r'; furthermore, if N is nonnegative the Z-transform of x converges also in the limit for $z \to \infty$. If x is left-sided, that is $x_n = 0$ for all $n > N$ and some $N \in \mathbb{Z}$, the region of convergence is a disc around the origin; for $N \leq 0$ the region of convergence includes the origin, $\mathrm{roc}(x) = \left\{z \in \mathbb{C}\colon |z| < r'\right\}$ for some real r'; for $N > 0$ the origin is excluded, $\mathrm{roc}(x) = \left\{z \in \mathbb{C}\colon z \neq 0, |z| < r'\right\}$. For any other signal x the region of convergence of x has annular shape, $\mathrm{roc}(x) = \{z \in \mathbb{C}\colon r_1 < |z| < r_2\}$ for some real r_1, r_2; the

[5] A precise definition of the spectral power density of a noise signal requires the machinery of stochastic processes.

Table 2.5 Some discrete-time signals and their Z-transforms

w_n	$\hat{W}(z) = (\mathcal{Z}(w))(z)$	$\mathrm{roc}(w)$
$ax_n + by_n$	$a\hat{X}(z) + b\hat{Y}(z)$	$\mathrm{roc}(w) \supseteq \mathrm{roc}(x) \cap \mathrm{roc}(y)$
x_{n-M}	$z^{-M}\hat{X}(z)$	$\mathrm{roc}(x)$
$(x * y)_n$	$\hat{X}(z)\hat{Y}(z)$	$\mathrm{roc}(w) \supseteq \mathrm{roc}(x) \cap \mathrm{roc}(y)$
x_n^*	$\hat{X}(z^*)^*$	$\mathrm{roc}(x)$
x_{-n}	$\hat{X}(z^{-1})$	$\left\{z\colon z^{-1} \in \mathrm{roc}(x)\right\}$
nx_n	$-z\dfrac{\mathrm{d}\hat{X}(z)}{\mathrm{d}z}$	$\mathrm{roc}(x)$
$a^{-n}x_n$	$\hat{X}(az)$	$\{z\colon az \in \mathrm{roc}(x)\}$
δ_{n-M}	z^{-M}	$\mathbb{C}$
$u_n = \begin{cases} 0 & \text{if } n < 0 \\ 1 & \text{if } n \geq 0 \end{cases}$	$\dfrac{z}{z-1}$	$\{z \in \mathbb{C}\colon \lvert z\rvert > 1\}$
$a^n u_n$	$\dfrac{z}{z-a}$	$\{z \in \mathbb{C}\colon \lvert z\rvert > \lvert a\rvert\}$
$a^n u_{-n}$	$\dfrac{a}{a-z}$	$\{z \in \mathbb{C}\colon \lvert z\rvert < \lvert a\rvert\}$

The signals $x\colon \mathbb{Z} \to \mathbb{C}$ and $y\colon \mathbb{Z} \to \mathbb{C}$ are arbitrary; $\hat{X}\colon \mathrm{roc}(x) \to \mathbb{C}$ and $\hat{Y}\colon \mathrm{roc}(y) \to \mathbb{C}$ are their Z-transforms. The numbers $M \in \mathbb{Z}$ and $a, b \in \mathbb{C}$ are arbitrary

region of convergence may also be empty in this case. The region of convergence is an integral part of the Z-transform; the Z-transform of some entirely different signals differ in the respective regions of convergence only (Table 2.5).

Provided the unit circle is part of the region of convergence of $\hat{X}$ the Z-transform evaluated on the unit circle yields the discrete-time Fourier transform of x,

$$X(\omega) = (\mathcal{F}(x))(\omega) = \hat{X}(\mathrm{e}^{\mathrm{i}\omega}).$$

2.7.1 Stability of Linear Time-Invariant Discrete-Time Systems

A linear time-invariant discrete-time system S is bounded-input bounded-output stable if and only if for all bounded inputs $x\colon \mathbb{Z} \to \mathbb{R}$ the output $y\colon \mathbb{Z} \to \mathbb{R}$, $n \mapsto y_n = (S(x))_n$ is also bounded. A signal $x\colon \mathbb{Z} \to \mathbb{C}$ is bounded if and only if $|x_n| < A$ for all $n \in \mathbb{Z}$ and for some $A \in \mathbb{R}$. A necessary and sufficient condition for stability is that the impulse response h of S is absolutely summable, that is $\sum_{n=-\infty}^{\infty} |h_n|$ converges to a finite value. Moreover, the system S is stable if and only if the unit circle is part of the region of convergence of the transfer function $\hat{H}$ of S, which is the Z-transform of the impulse response h of S.

The transfer function of many linear time-invariant discrete-time systems is the quotient of two polynomials, $\hat{H}(z) = \frac{A(z)}{B(z)}$. The zeros of the numerator A are called the zeros of $\hat{H}$; the zeros of the denominator B are the poles of $\hat{H}$. A causal linear time-invariant discrete-time system S is stable if and only if all poles of the transfer function $\hat{H}$ of S lie within the unit circle.

Using the Z-transform we can argue about the stability and the frequency response of a composition of discrete-time systems in which some of the constituents are instable.

2.8 The Two-Sided Laplace Transform

Let $x\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto x(t)$ be a continuous-time signal. The two-sided Laplace transform[6] of x, $\hat{X}\colon \mathrm{roc}(x) \to \mathbb{C}$ is defined as

$$s \mapsto \hat{X}(s) = (\mathcal{L}(x))(s) = \int_{-\infty}^{\infty} x(t)\mathrm{e}^{-st}\mathrm{d}t. \tag{2.22}$$

The set $\mathrm{roc}(x)$ is the region of convergence of the Laplace transform of x. When we substitute $s = \sigma + \mathrm{i}\omega$ with $\sigma, \omega \in \mathbb{R}$ in (2.22) we recognize that the Laplace transform is actually the continuous Fourier transform of the signal $x'\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto x'(t) = x(t)\mathrm{e}^{-\sigma t}$. If the signal x' satisfies the Dirichlet conditions the integral in (2.22) converges. Assuming the signal x is piecewise continuous and of bounded variation in every finite interval—assumptions satisfied by the signals appearing in practice—then the integral in (2.22) converges provided $\int_{-\infty}^{\infty} |x(t)\mathrm{e}^{-\sigma t}|\mathrm{d}t$ converges to a finite value.

If x is right-sided, that is $x(t) = 0$ for all $t < T$ and some $T \in \mathbb{R}$, the region of convergence is $\mathrm{roc}(x) = \left\{s \in \mathbb{C}\colon \Re(s) > \sigma'\right\}$ for some real σ'. If x is left-sided, that is $x(t) = 0$ for all $t > T$ and some $T \in \mathbb{R}$, the region of convergence is $\mathrm{roc}(x) = \left\{s \in \mathbb{C}\colon \Re(s) < \sigma'\right\}$ for some real σ'. For any other signal x the region of convergence of x is $\mathrm{roc}(x) = \{s \in \mathbb{C}\colon \sigma_1 < \Re(s) < \sigma_2\}$ for some real σ_1, σ_2; the region of convergence may also be empty in this case. The region of convergence is an integral part of the Laplace transform; the Laplace transform of some entirely different signals differ in the respective regions of convergence only (Table 2.6).

Let $x\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto x(t)$ be a signal with $x(t) = 0$ for $t < 0$; let the signal $X\colon \mathbb{C} \to \mathbb{C}$ be the Laplace transform of x. Let the imaginary axis be part of $\mathrm{roc}(x)$.

---

[6] In the one-sided Laplace transform integration ranges from 0 to ∞. The one-sided transform is useful for solving initial value problems involving linear differential equations. See also Appendix C.

Table 2.6 Some continuous-time signals and their Laplace transforms

$w(t)$	$\hat{W}(s) = (\mathcal{L}(w))(s)$	$\mathrm{roc}(w)$
$ax(t) + by(t)$	$a\hat{X}(s) + b\hat{Y}(s)$	$\mathrm{roc}(w) \supseteq \mathrm{roc}(x) \cap \mathrm{roc}(y)$
$x(t-\tau)$	$\mathrm{e}^{-s\tau}\hat{X}(S)$	$\mathrm{roc}(x)$
$(x * y)(t)$	$\hat{X}(s)\hat{Y}(s)$	$\mathrm{roc}(w) \supseteq \mathrm{roc}(x) \cap \mathrm{roc}(y)$
$x^*(t)$	$\hat{X}(s^*)^*$	$\mathrm{roc}(x)$
$x(ct)$	$\dfrac{\hat{X}(s/c)}{\lvert c\rvert}$	$\{s \in \mathbb{C} \colon s/c \in \mathrm{roc}(x)\}$
$tx(t)$	$-\dfrac{\mathrm{d}\hat{X}(s)}{\mathrm{d}s}$	$\mathrm{roc}(x)$
$\mathrm{e}^{at}x(t)$	$\hat{X}(s-a)$	$\{s \in \mathbb{C} \colon s - a \in \mathrm{roc}(x)\}$
$\int_{-\infty}^{t} x(\tau)\mathrm{d}\tau$	$\dfrac{\hat{X}(s)}{s}$	$\mathrm{roc}(w) \supseteq \{s \in \mathrm{roc}(x) \colon \Re(s) > 0\}$
$\dfrac{\mathrm{d}x(t)}{\mathrm{d}t}$	$s\hat{X}(s)$	$\mathrm{roc}(w) \supseteq \mathrm{roc}(x)$
$\delta(t-\tau)$	$\mathrm{e}^{-s\tau}$	$\mathbb{C}$
$u(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0 \end{cases}$	$\dfrac{1}{s}$	$\{s \in \mathbb{C} \colon \Re(s) > 0\}$
$\mathrm{e}^{-at}u(t)$	$\dfrac{1}{s+a}$	$\{s \in \mathbb{C} \colon \Re(s) > -\Re(a)\}$
$-\mathrm{e}^{-at}u(-t)$	$\dfrac{1}{s+a}$	$\{s \in \mathbb{C} \colon \Re(s) < -\Re(a)\}$

The signals $x \colon \mathbb{R} \to \mathbb{C}$ and $y \colon \mathbb{R} \to \mathbb{C}$ are arbitrary; $\hat{X} \colon \mathrm{roc}(x) \to \mathbb{C}$ and $\hat{Y} \colon \mathrm{roc}(y) \to \mathbb{C}$ are their Laplace transforms. The numbers $a, b \in \mathbb{C}$ and $c, \tau \in \mathbb{R}$ are arbitrary

The final value theorem then allows us to compute the limit of $f(t)$ for $t \to \infty$ from the Laplace transform X,

$$\lim_{t\to\infty} x(t) = \lim_{s\to 0} sX(s). \tag{2.23}$$

Provided the imaginary axis is part of the region of convergence of x evaluating the Laplace transform along the imaginary axis yields the continuous-time Fourier transform of x,

$$X(\omega) = (\mathcal{F}(x))(\omega) = \hat{X}(\mathrm{i}\omega).$$

2.8.1 Stability of Linear Time-Invariant Continuous-Time Systems

A linear time-invariant continuous-time system S is bounded-input bounded-output stable if and only if for all bounded inputs $x\colon \mathbb{R} \to \mathbb{R}$ the output $y\colon \mathbb{R} \to \mathbb{R}$, $n \mapsto y(t) = (S(x))(t)$ is also bounded. A signal $x\colon \mathbb{R} \to \mathbb{R}$ is bounded if and only if $|x(t)| < A$ for all $t \in \mathbb{R}$ and for some $A \in \mathbb{R}$. A necessary and sufficient condition for stability is that the impulse response h of S is absolutely integrable, that $\int_{-\infty}^{\infty} |h(t)| \mathrm{d}t$ converges to a finite value.

The transfer function $\hat{H}$ of the system S is the Laplace transform of the impulse response of S, $\hat{H}\colon \mathbb{C} \to \mathbb{C}$,

$$s \mapsto \hat{H}(s) = \frac{\hat{Y}(s)}{\hat{X}(s)} = \mathcal{L}(h)(s),$$

where $\hat{X}$ is the Laplace transform of an input x and $\hat{Y}$ is the Laplace transform of the corresponding output y. The system S is stable provided the imaginary axis is part of the region of convergence of the transfer function $\hat{H}$ of S.

The transfer function of many linear time-invariant continuous-time systems is the quotient of two polynomials, $\hat{H}(z) = \frac{A(z)}{B(z)}$. The zeros of the numerator A are called the zeros of $\hat{H}$; the zeros of the denominator B are the poles of $\hat{H}$. A causal linear time-invariant continuous-time system S is stable provided all poles of the transfer function $\hat{H}$ of S have a negative real part.

Using the Laplace transform, we can argue about the stability and the frequency response of a composition of continuous-time systems in which some of the constituents are instable.

Let us consider the continuous-time system S described by the equation

$$y(t) = x(t) - \frac{1}{\beta} \int_{-\infty}^{t} y(\tau) \mathrm{d}\tau, \tag{2.24}$$

for $\beta > 0$. Applying the Laplace transform to both sides and rearranging we get

$$\frac{\hat{Y}(s)}{\hat{X}(s)} = \hat{H}(s) = \frac{s\beta}{1 + s\beta}.$$

There are two possibilities for the region of convergence of the impulse response h of S, $\mathrm{roc}(h) = \{s \in \mathbb{C}\colon \Re(s) < -1/\beta\}$ describing a noncausal system and $\mathrm{roc}(h) = \{s \in \mathbb{C}\colon \Re(s) > -1/\beta\}$ describing a stable causal system. The frequency response of S for the second choice is

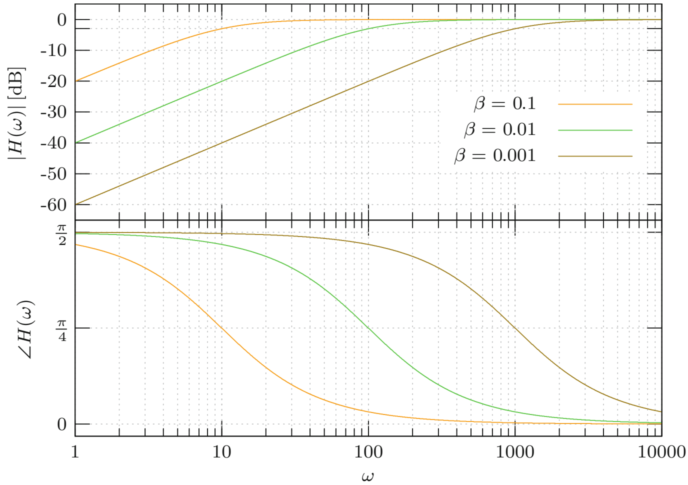


Fig. 2.2 Bode plot of the continuous-time system described by Eq. (2.24)

$$H(\omega) = \hat{H}(\mathrm{i}\omega) = \frac{\omega\beta}{\omega\beta - \mathrm{i}}.$$

We plot the magnitude and the phase of the frequency response separately, Fig. 2.2. For the magnitude we use the decibel scale. For the angular frequency ω we choose a logarithmic scale. Inspecting this so-called Bode plot we recognize the system S as a highpass filter.

2.9 Differential Equations, State-Space Models

The system of n linear differential equations with constant coefficients

$$\begin{aligned}
\frac{\mathrm{d}v_1(t)}{\mathrm{d}t} &= a_{1,1}v_1(t) + a_{1,2}v_2(t) + \cdots + a_{1,n}v_n(t) + b_1x(t)\\
\frac{\mathrm{d}v_2(t)}{\mathrm{d}t} &= a_{2,1}v_1(t) + a_{2,2}v_2(t) + \cdots + a_{2,n}v_n(t) + b_2x(t)\\
&\vdots\\
\frac{\mathrm{d}v_n(t)}{\mathrm{d}t} &= a_{n,1}v_1(t) + a_{n,2}v_2(t) + \cdots + a_{n,n}v_n(t) + b_nx(t)\\
y(t) &= c_1v_1(t) + c_2v_2(t) + \cdots + c_nv_n(t) + dx(t),
\end{aligned}$$

where $a_{i,j} \in \mathbb{R}$, $b_i \in \mathbb{R}$, $c_i \in \mathbb{R}$ and $d \in \mathbb{R}$ for $1 \leq i, j \leq n$, constitute the state-space model of a linear time-invariant continuous-time system S with one input $x(t)$ and one output $y(t)$. The signals $v_1(t), \ldots, v_n(t)$ are the states. For simulating the system S with tools like the Simulink® system we may either use the state space model directly or translate it into a block diagram. In order to construct such a block diagram we introduce for each state $v_i(t)$ in the state-space model an integrator into the block diagram. The output of the integrator produces the state $v_i(t)$, while the input of the integrator is wired to a block diagram computing the right-hand side of the equation for $\frac{dv_i(t)}{dt}$ in the state-space model. The outputs of the integrators are fed back into the right-hand sides. Another block diagram computes the output from the input and the states. For an example see Fig. 2.6.

The state space representation is not limited to linear systems. As long as we manage to transform the system of differential equations into equations defining the n states of the form

$$\frac{dv_i(t)}{dt} = f_i(v_1(t), v_2(t), \ldots, v_n(t), x(t))$$

and an equation defining the output of the form

$$y(t) = g(v_1(t), v_2(t), \ldots, v_n(t), x(t))$$

we have a state space model suitable for simulation.

2.9.1 The Harmonic Oscillator

The apparatus of differential equations was invented for modeling physical systems. These systems can be mechanical, electrical, optical or any other domain imaginable. Let us, for example, consider a body with mass m suspended from the ceiling via a spring, Fig. 2.3. We assume the spring has no mass and that the spring follows Hook's law. More precisely, we assume that the force $F_S(t)$ the spring exerts on the body is proportional to the displacement $y(t)$ of the body from rest, $F_S(t) = -ky(t)$ where k is the stiffness of the spring. Measurements show that Hook's law describes the action of a spring made from steel reasonably well. The body is attached to a damper, which is fixed to the floor. We assume the damper exerts a force on the body proportional to the body's velocity, $F_D(t) = -c\frac{dy(t)}{dt}$. The constant of proportionality c is called

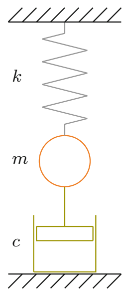


Fig. 2.3 A body is suspended from the ceiling with a spring. The body is connected to a damper, which is fixed to the floor

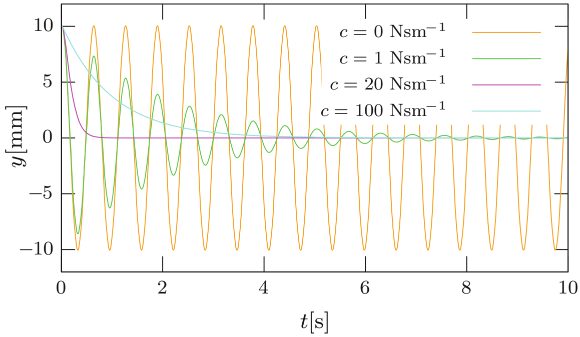


Fig. 2.4 Behavior of a mass-spring-damper system for different damping coefficients. The body has a mass of 1 kg, the spring a stiffness of 100 Nm^{-1}. The damping coefficient $c = 20$ Nsm^{-1} results in the system being critically damped

the damping constant. No external forces shall act on our contraption. According to Newton's second law of motion, the sum of these two forces have to balance the force F_m required for accelerating the body, $F_m(t) = m\frac{\mathrm{d}^2y(t)}{\mathrm{d}t^2}$. Equating $F_m(t) = F_S(t) + F_D(t)$ and rearranging gives us the second order differential equation

$$m\frac{\mathrm{d}^2y(t)}{\mathrm{d}t^2} + c\frac{\mathrm{d}y(t)}{\mathrm{d}t} + ky(t) = 0, \tag{2.25}$$

the equation of a harmonic oscillation.

Depending on the magnitude of the damping constant c, we can identify four different types of solutions for (2.25), Fig. 2.4. Let $\alpha = \frac{c}{2m}$. If $c = 0$ the general solution has the form

$$y(t) = A\cos\omega_0 t + B\sin\omega_0 t,$$

where $\omega_0 = \sqrt{\frac{k}{m}}$, which describes an undamped harmonic oscillation. For $c^2 < 4mk$ the general solution is

$$y(t) = \mathrm{e}^{-\alpha t}\left(A\cos\omega' t + B\sin\omega' t\right),$$

where $\omega' = \sqrt{\frac{k}{m} - \frac{c^2}{4m^2}}$, which describes a damped harmonic oscillation. For $c^2 = 4mk$ the general solution is

$$y(t) = (A + Bt)\mathrm{e}^{-\alpha t},$$

which describes a critically damped system. A critically damped system will return to rest without oscillation as fast as possible. At last, for $c^2 > 4mk$ the general solution is

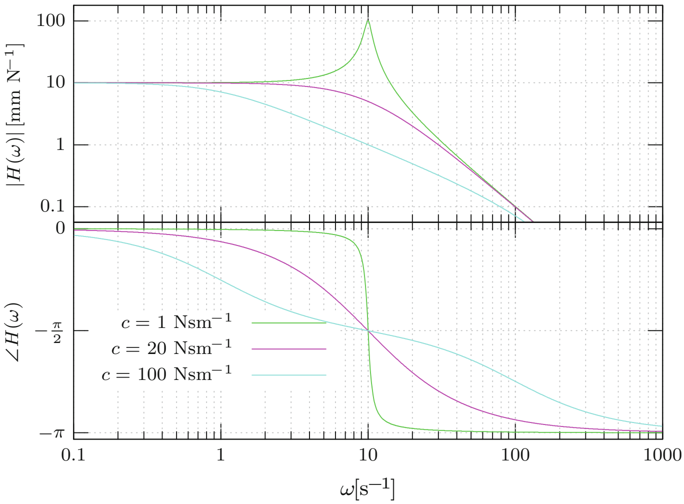


Fig. 2.5 Frequency response of the driven harmonic oscillator for different damping coefficients. The damping coefficient $c = 1$ Nsm^{-1} results in a peak in the amplitude response indicating resonance at that frequency

$$y(t) = A\mathrm{e}^{-(\alpha-\beta)t} + B\mathrm{e}^{-(\alpha+\beta)t},$$

where $\beta = \sqrt{\frac{c^2}{4m^2} - \frac{k}{m}}$, which describes an overdamped system.

In order to handle an external driving force $x(t)$ acting on the body we have to change (2.25) into

$$m\frac{\mathrm{d}^2 y(t)}{\mathrm{d}t^2} + c\frac{\mathrm{d}y(t)}{\mathrm{d}t} + ky(t) = x(t), \tag{2.26}$$

the equation of a driven harmonic oscillator. For analyzing the behavior of this system, we compute the system's transfer function $\hat{H}(s)$ by applying the Laplace transform to both sides of (2.26) and rearranging,

$$\hat{H}(s) = \frac{\hat{Y}(s)}{\hat{X}(s)} = \frac{1}{ms^2 + cs + k}.$$

The frequency response $H(\omega) = \hat{H}(\mathrm{i}\omega)$, Fig. 2.5, shows a peak of the amplitude at

$$\omega_\mathrm{r} = \sqrt{\frac{k}{m} - \frac{c^2}{2m^2}}$$

for $c^2 < 2mk$. When the external force excites the system close to that frequency, the amplitude of the response will be much larger than for external forces at other frequencies. This behavior is called resonance; $f_\mathrm{r} = \frac{\omega_\mathrm{r}}{2\pi}$ is the resonant frequency

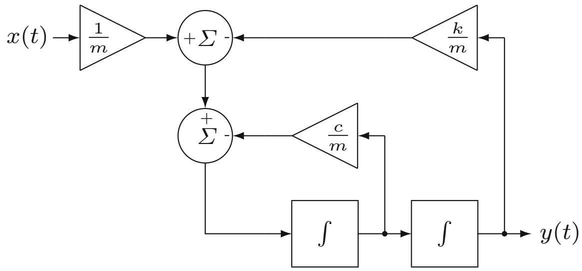


Fig. 2.6 Block diagram for Eq. (2.26)

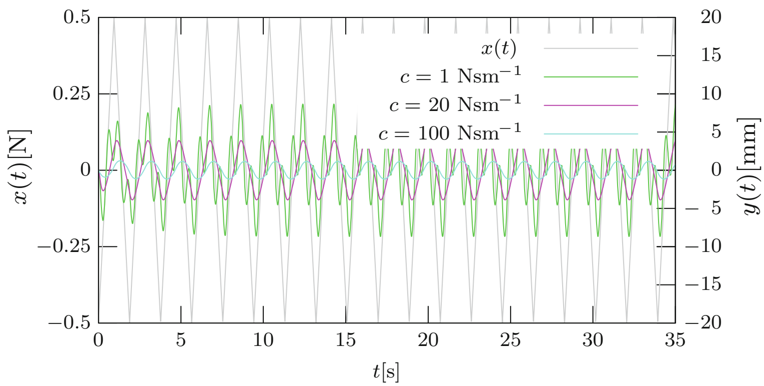


Fig. 2.7 Numerical simulation of the mass-spring-damper system. The underdamped system shows resonance at the third harmonic of the driving force, while the critically damped and the overdamped system show no such reaction

of the driven mass-spring-damper system. Unexpected resonance can be disastrous; the large excursions associated with it can overload structures.

Alternatively, we can bring (2.26) into state-space form,

$$
\begin{aligned}
\frac{\mathrm{d}v_1(t)}{\mathrm{d}t} &= v_2(t) \\
\frac{\mathrm{d}v_2(t)}{\mathrm{d}t} &= -\frac{c}{m}v_2(t) - \frac{k}{m}v_1(t) + \frac{1}{m}x(t) \\
y(t) &= v_1(t),
\end{aligned}
$$

and transform it into a block diagram for numerical simulation, Fig. 2.6. When we drive the system with a periodic, triangular force with about one third of the frequency f_r, then the content of the driving force at the third harmonic drives the underdamped system into resonance, Fig. 2.7. The driving force sets in at time 0, while it was zero before. The critically damped system and the overdamped system show no recognizable reaction to the excitation at the third harmonic. The simulations in Fig. 2.7 start out with the systems at rest. Each of the three systems takes some time for reaching a steady response. The output of a linear time-invariant system S to a sinusoidal input $x(t)$, which sets at time 0, $x(t) = u(t)a\cos(\omega t + \phi)$, is the sum of the so-called transient response, which for a stable system S tapers off after some time, and the steady state response. The steady state response has the same frequency as the input. A frequency response-based analysis of the system S provides us with the steady state information only.

2.9.2 The Pendulum

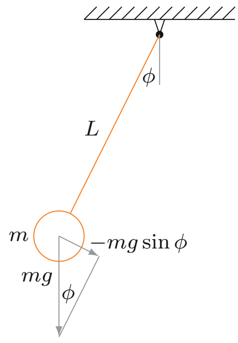
An ideal pendulum, Fig. 2.8, consists of a massless stiff rod of length L and a point-shaped body with mass m. One end of the rod attaches to the mass, while the other connects to a joint. The joint allows the rod together with the body to rotate freely on a horizontal axis. Gravity pulls with a force of mg at the mass. When the rod is at an angle $\phi(t)$ toward the vertical the force $F_g = -mg\sin\phi(t)$ tries to restore the pendulum to the vertical position. The constant $g = 9.81\,\mathrm{ms}^{-2}$ is the Earth's gravitational pull. The force F_g has to balance the force $F_m = mL\frac{\mathrm{d}^2\phi(t)}{\mathrm{d}t^2}$ for the tangential acceleration of the pendulum, $F_m = F_g$. Rearranging yields

$$mL\frac{\mathrm{d}^2\phi(t)}{\mathrm{d}t^2} + mg\sin(\phi(t)) = 0.$$

This differential equation is nonlinear. We introduce a damping force F_D, which is proportional to the velocity of the mass, $F_D = -Lc\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}$, then $F_m = F_D + F_g$ and

$$\frac{\mathrm{d}^2\phi(t)}{\mathrm{d}t^2} + \frac{c}{m}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} + \frac{g}{L}\sin(\phi(t)) = 0.$$

Fig. 2.8 A body with mass m connected to a rod, a pendulum, is allowed to pivot on a horizontal axle. Whenever the pendulum deviates from the vertical the Earth's gravitation provides a restoring force



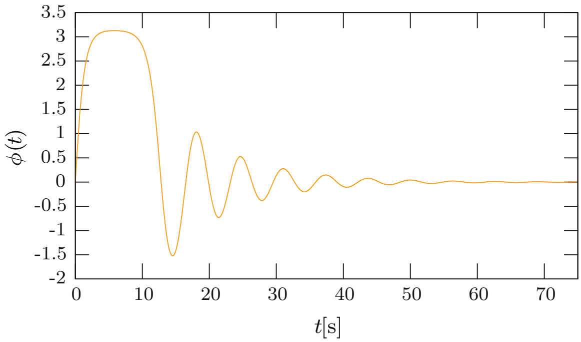


Fig. 2.9 Behavior of a pendulum starting with an initial angular velocity. The length of the rod is $L = 10\,\text{m}$, the mass is 1 kg and the damping constant is $c = 0.2\ \text{Nsm}^{-1}$

We can convert this equation into state-space form and simulate the pendulum's behavior numerically, Fig. 2.9. Note that the time for one oscillation depends on the amplitude of the oscillation. The nonlinear nature of the driven pendulum puts its analysis beyond the means of the mathematical tools we discuss.

2.10 Feedback Control

Often we want the output $y(t)$ of a causal system S_1 to track a reference signal $r(t)$. More precisely, we want to find an input $u(t)$ for the system S_1 such that $(S_1(u))(t)$ is as close to $r(t)$ as possible. Control engineers study this problem extensively. The system S_1, called the plant, consists of an actuator for converting the input $u(t)$ into some physical quantity—such as the position of a valve or the speed of a motor—and the process producing the output.

Provided we know the plant S_1 exactly we may try to find a causal system S_2, a controller, such that the series composition of the controller S_2 with the plant S_1 tracks the reference $r(t)$ reasonably well, in other words that $(S_1(S_2(r)))(t)$ is reasonably close to $r(t)$. The transfer function of this series composition is $\hat{G}_2(s)\hat{G}_1(s)$, where $\hat{G}_1$ is the transfer function of the plant S_1 and $\hat{G}_2$ is the transfer function of the controller S_2. In this open-loop operation the controller cannot react to disturbances in the plant as it lacks information about the output $y(t)$. More severely, open-loop operation does not work when the plant S_1 is unstable. While we may think that we can cancel unwanted poles p with $\Re(p) \geq 0$ in the transfer function $\hat{G}_1(s)$ of the plant S_1 with corresponding zeros of $\hat{G}_2(s)$, deteriorating conditions over the lifetime of the plant and the controller will move poles and zeros rendering such an attempt futile.

The physics of the plant S_1 determines whether the plant is stable or not. For controlling a possibly instable plant, which is subjected to disturbances, we have to feed back information about the output $y(t)$ to the controller S_2. In the unity feedback composition, Fig. 2.10 left, the output $y(t)$ is subtracted from the reference $r(t)$ producing the error signal $e(t)$. The controller S_2 tries to minimize the error signal by producing the input $u(t)$ for the plant S_1 according to an appropriate control law. Feeding back the output to the controller creates a closed control loop; subtracting the output from the reference makes the feedback negative. The unity feedback composition has the transfer function

$$\hat{H}(s) = \frac{\hat{G}_1(s)\hat{G}_2(s)}{1 + \hat{G}_1(s)\hat{G}_2(s)},$$

where $\hat{G}_1(s)$ is the transfer function of the plant and $\hat{G}_2(s)$ is the transfer function of the controller, while the feedback composition in Fig. 2.10 right, has the transfer function

$$\hat{H}'(s) = \frac{\hat{G}_1(s)}{1 + \hat{G}_1(s)\hat{G}_2(s)}.$$

The latter composition appears in the analysis of circuits involving operational amplifiers. An operational amplifier is an electronic amplifier, which amplifies the difference between its two inputs, ideally with infinite gain. The operational amplifier then acts as plant, while the so-called feedback network acts as controller.

In both feedback compositions the poles of the plant's transfer function $\hat{G}_1(s)$ cancel. By designing an appropriate controller S_2 we can place the poles and zeros of the chosen feedback composition structure as we like.

Sometimes we do not know the transfer function of the plant, but we still can measure the frequency response $H(\omega)$ of the plant S_1, provided the plant is stable. We consider a proportional controller with the transfer function $\hat{G}_2(s) = K_\mathrm{P}$ for some gain $K_\mathrm{P} \in \mathbb{R}$, $K_\mathrm{P} > 0$ and assume that the unity feedback system built from S_1 and S_2 is stable for small gains K_P and becomes unstable for large gains. The phase margin then is $\angle H(\omega') + \pi$ where $|K_\mathrm{P} H(\omega')| = 1$ for $\omega' \in \mathbb{R}$, $\omega' > 0$. If this phase margin is positive then the unity feedback composition of S_1 and S_2 is stable. A controller introducing delay, due to a software implementation for example, eats up some of the phase margin designed into the control system, in that way

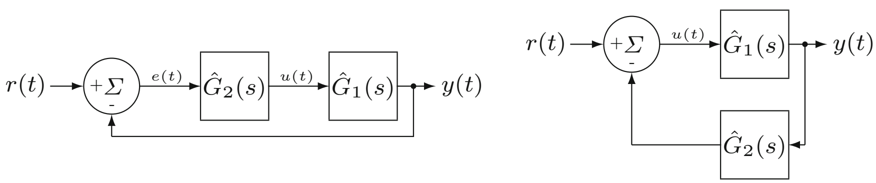


Fig. 2.10 Two possible feedback compositions of one system, the plant, with transfer function $\hat{G}_1(s)$ and one, the controller, with transfer function $\hat{G}_2(s)$. The unity feedback composition to the *left* is used for building control systems, while the composition to the *right* appears in the analysis of circuits involving operational amplifiers

diminishing the quality of the control. Therefore, unnecessary delays in a controller's implementation shall be avoided.

A proportional controller employed in the unity feedback configuration feeds a scaled version of the error signal to the plant, $u(t) = K_P e(t)$. We look at the response of the control loop to a unit-step input. The closed-loop transfer function is

$$\hat{H}(s) = \frac{K_P \hat{G}_1(s)}{1 + K_P \hat{G}_1(s)}.$$

The Laplace transform of the step response is $\hat{H}(s)/s$. Assuming stability of the closed-loop system and using the final value theorem (2.23) we compute the steady state,

$$\lim_{t\to\infty} y(t) = \lim_{s\to 0} s\frac{\hat{H}(s)}{s} = \hat{H}(0) = \lim_{s\to 0} \frac{K_P \hat{G}_1(s)}{1 + K_P \hat{G}_1(s)}.$$

The steady state response is unequal to one whenever $\lim_{s\to 0} \hat{G}_1(s)$ is finite. For making the steady state error $|1 - \hat{H}(0)|$ disappear the transfer function $\hat{G}_1(s)$ of the plant must have a pole at 0. Increasing the controller's gain K_P will reduce the steady state error at the cost of increasing the overshoot of the output $y(t)$ whenever the reference $r(t)$ changes.

In order to make the steady state error vanish for plants having no pole at 0 we augment the proportional controller by an integrator resulting in a proportional plus integral controller, its control equation is

$$u(t) = K_P \left(e(t) + \frac{1}{T_I} \int_{t_0}^{t} e(\tau) \mathrm{d}\tau \right), \tag{2.27}$$

where t_0 is the time the control loop starts to operate. All signals associated with the control loop are assumed to be zero before t_0. The controller's transfer function $\hat{G}_2(s)$ is

$$\hat{G}_2(s) = K_P \left(1 + \frac{1}{T_I s}\right).$$

As rule of thumb, increasing the proportional gain K_P makes the control loop react faster to changes of the reference at the cost of increased overshoot. It also reduces the steady state error. Decreasing the integral time T_I speeds up the control loop too, again at the cost of increased overshoot.

Real plants have physical limits beyond which they cannot operate. A valve for example cannot operate beyond fully opened or fully closed. When such operational limits prevent the plant's output to reach the reference, the integrator's output will either grow or shrink without bounds. When the reference later returns to a reachable value it will take the integrator considerable time for winding down to sensible values.

Fig. 2.11 Basic structure of the software realization of a proportional and integral controller

procedure PICONTROLLER(y_k,r_k)
  $e_k \leftarrow r_k - y_k$
  $u_k \leftarrow u_{k-1} + K_\mathrm{P}\left(e_k + \frac{T}{T_\mathrm{I}} e_k - e_{k-1}\right)$
  output u_k
  $e_{k-1} \leftarrow e_k$
  $u_{k-1} \leftarrow u_k$
end procedure

During this time the input to the plant will still exceed the plant's limits. A crude anti-windup strategy limits the value of the integrator's output by temporarily switching off the integrator.

The procedure PICONTROLLER in Fig. 2.11 is the implementation of a proportional and integral controller in software. It must be called periodically with period T. We arrive at this discrete-time realization by approximating the integral in (2.27) with a sum. The procedure accepts the plant's output measured for the kth invocation y_k and the reference r_k as parameters. The error e_{k-1} and the output u_{k-1}, both from the previous invocation, constitute the state of the controller. It is good practice to pass the plant's new input before updating the controller's state, in order to avoid unnecessary delays.

For reducing the overshoot we may add a differential term to the proportional plus integral controller resulting in a proportional integral differential controller,

$$u(t) = K_\mathrm{P}\left(e(t) + \frac{1}{T_\mathrm{I}}\int_{t_0}^{t} e(\tau)\mathrm{d}\tau + T_\mathrm{D}\frac{\mathrm{d}e(t)}{\mathrm{d}t}\right).$$

As a rule of thumb increasing the derivative time T_D reduces the overshoot. The differential term, however, is sensitive to noise added when measuring the plant's output $y(t)$.

In many control systems several control loops are nested into each other. The outermost control loop in such a cascaded control scheme processes the reference $r(t)$. The output of the controller in the outer loop is the reference for the next inner control loop. When controlling an electric motor, for example, the innermost loop controls the torque the motor produces. The next outer loop controls the velocity by passing the torque required for achieving the desired velocity as reference to the innermost loop. In drives used for positioning, the outermost loop controls the position by passing the velocity required for arriving and holding the desired position to the middle control loop.

Fig. 2.12 A function generator produces periodic signals within a frequency range from $1\,\mu$Hz up to several megahertz

2.11 Instrumentation for Producing and Measuring Signals

When we engineer an embedded system our ideas eventually must prove themselves in the physical reality. In order to quantify the performance of our system we must stimulate it with well-defined inputs, and measure and analyze its responses. We limit our discussion to basic instrumentation for general purpose use.

2.11.1 The Function Generator

A function generator[7] produces periodic signals with frequencies ranging from zero to several megahertz, Fig. 2.12. It offers the choice between several predefined waveforms; sine, rectangle, ramp, triangle, pulse, and noise are typical. A freely programmable waveform may also be available. Frequency, amplitude, and a constant offset from zero can be set for the selected waveform. For the rectangle waveform we can select the so-called duty cycle, that is the ratio between the time the signal is on and the time the signal is off. For the pulse waveform we can set the width of the pulses. The noise waveform typically is band-limited white noise.

In addition, a function generator offers basic modulation features. Modulation means that the output signal results from the modification of a higher frequency carrier signal with a lower frequency signal. Amplitude modulation multiplies the

[7] The ancestor of today's function generators is the HP-200A, an audio oscillator, which grew out of Bill Hewlett's master's thesis at Stanford. The HP-200A is famous for the ingenious use of a light bulb for providing negative feedback. Bill Hewlett and Dave Packard assembled these oscillators starting 1939 in the garage behind Packard's house at 367 Addison Avenue in Palo Alto. One of the first customers, buying eight HP-200B, were the Walt Disney Studios.

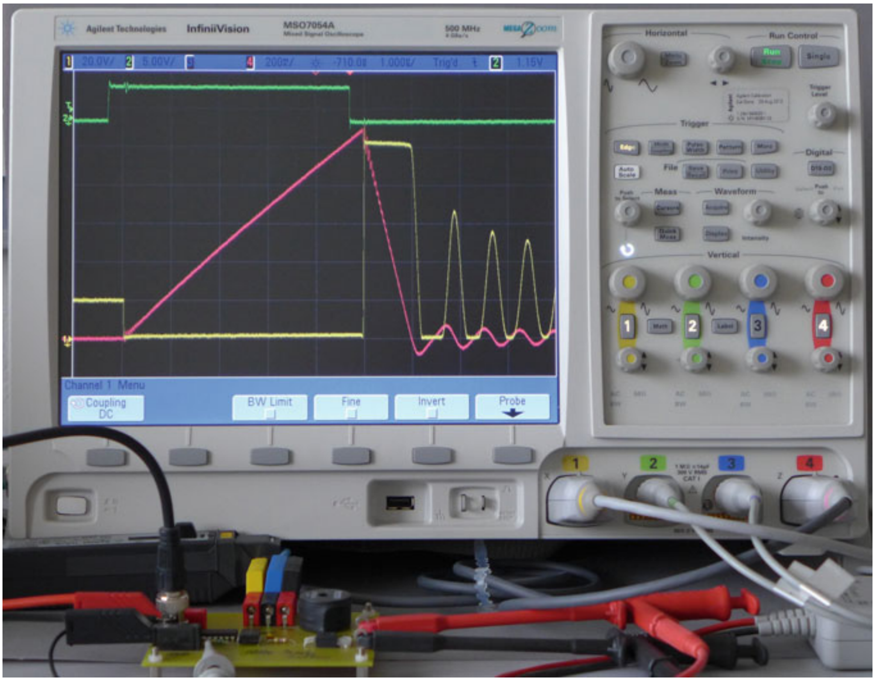


Fig. 2.13 Still life with oscilloscope. An oscilloscope simultaneously records and displays several signals. The length of a record can vary between less than a nanosecond and several hundred seconds. The range of possible amplitudes can span more than three decades

carrier with the lower frequency signal for producing the output. Frequency modulation shifts the frequency of the carrier according to the lower frequency signal.

A frequency sweep modulates the frequency of the output signal with a slow ramp. By analyzing the answer of some device to such a frequency sweep we can derive the device's frequency response. A linear sweep uses a sawtooth-shaped ramp. A logarithmic sweep uses an exponential ramp as modulation. For measurement instruments analyzing signals derived from the function generator's output the function generator provides a synchronization signal indicating, for example, the beginning of a sweep.

2.11.2 The Oscilloscope

The oscilloscope, scope for short, is the most important instrument for capturing and displaying signals, Fig. 2.13. A scope consists of a horizontal section, a vertical section, a trigger section[8] and a display for the waveforms. Each section has its own set of controls on the front panel. Modern scopes usually have several channels, two or four are common, for simultaneously capturing and displaying several signals.

The vertical section consists of the amplifiers for the channels. The controls for each channel are the gain, the channel's vertical position on the scope's screen, the coupling, and a switch for disabling the channel. A channel's gain is usually shown in a line above the screen's graticule. The unit the gain is stated in is the unit of the measured physical quantity per vertical division of the graticule. The control for the vertical position allows us to arrange the channels on the screen. A small triangle to the left of the graticule indicates zero for the channel. The coupling can be alternating current (AC) or direct current (DC). With the coupling set to AC a highpass filter, which blocks constant signal content, is introduced right after the channel's input connector. With DC coupling constant signal content is passed to the channel's amplifier.

When the trigger section recognizes a trigger event it commands the scope to start a new recording. The most basic type of trigger event is when the signal fed into one of the scope's channels crosses an adjustable trigger level, either from above or from below. A small triangle marked with the letter T to the left of the graticule indicates the trigger level. The operator can select between AC and DC coupling for the trigger source. Sophisticated trigger sections support triggers, for example, on pulse width, or on the appearance of some pattern on optional digital inputs. Practically, all scopes support an external trigger source.

The horizontal section provides the timing for the recordings. In analog scopes the horizontal section provided a ramp to the horizontal deflection coils of a cathode ray tube whenever a trigger section commanded a new recording. The horizontal coils deflected the electron beam inside the tube from left to right. Where the electron beam hit the phosphorus covering the tube's screen the phosphorus produced a spot of light. The output of the vertical amplifier drove the vertical coils, which deflected the beam in the vertical direction. The afterglow of the beam's trace provided a visual image of the signal's waveform. In modern digital scopes memory replaces the short-term storage action of the cathode ray tube. The outputs of the vertical amplifiers are digitized at a rate controlled by the horizontal section and written into the scope's acquisition memory. The size of the acquisition memory determines how long the scope can record after a trigger event. After the scope has displayed the content of the acquisition memory on its screen it is ready to make a new recording. The scope can operate either in single trigger mode or in continuous trigger mode. In single trigger mode, the first trigger event commands a recording and all subsequent trigger events are ignored. Single trigger mode helps analyzing nonrecurring events. In continuous trigger mode, the scope takes a record whenever the trigger section commands it, and the scope is ready. The scope displays subsequent recordings on top of each other. In the process it lets old recordings fade away.

---

[8] The basic concept of a scope has changed very little since Howard Vollum and Melvin Murdock, the founders of Tektronix, introduced the first practical oscilloscope, the Tektronix type 511, in 1948.

Digital scopes allow us to apply mathematical operations to the recordings in real time. The absolute value of the recording's Fourier transform is particularly useful for estimating the frequency content of a measured signal.

Scopes are not perfect either. High frequency signals are not reproduced faithfully. A scope's bandwidth specifies the frequency at which the displayed amplitude is down by 3dB from the true value. For observing digital waveforms with some accuracy the scope's bandwidth should be at least five times higher than the frequency of the fastest digital clock signal. For observing analog waveforms the scope's bandwidth shall be three times higher than the highest harmonic in the observed signals.

2.12 Bibliographical Notes

The book (Lee and Varaiya 2011) stands out for treating discrete-time and continuous-time equally. The book by Kreyszig (2010) provides a broad reference to many mathematical topics of use when doing engineering work. The book (Oppenheim and Schafer 1975) is still a very useful reference for discrete-time signals. Jänich's book (2001) covers advanced material such as partial differential equations in a very readable manner. The book by Gershenfeld (1999) contains an ample selection of mathematical methods for modeling physical systems.

The book (Middleton 1996) contains an in-depth treatise on signals influenced by noise. For the effects of noise on the discrete Fourier transform of a signal see (Schoukens and Renneboog 1986; Kester and Analog Devices 2003). The representation of a periodic signal by its Fourier series is proven for example in the book by Zygmund (2002a, Chap. 2, Theorem 8.1). The representation of a signal by its Fourier integral is proven for example in (Zygmund 2002b, Chap. 16, Theorem 1.3 and the text following it). For both continuous-time and discrete-time control systems see for example (Franklin et al. 2010, 1997). Both books contain many exemplary models of physical systems.

2.13 Exercises

Exercise 2.1 Several functions are used to suppress artifacts in the Fourier transform of functions which we have observed for a finite amount of time only. In this context these functions are called windows. The idea is to observe the signal $s\colon \mathbb{R} \to \mathbb{C}$, $t \mapsto s(t)$ only between times $-p$ and p for some positive real p and pad the observation $o(t)$ with zeros before and after. Let us consider the signal $s(t) = 1$, whose transform is the Dirac delta $\delta(\omega)$. The observation $o(t)$ is the appropriate boxcar function. The transform of the boxcar function shows us the errors we introduce into the Fourier transform by replacing the signal $s(t)$ by the observation $o(t)$. Plot the Fourier transform of a boxcar window.

Exercise 2.2 We can modify the error in the Fourier transform by transforming the signal $w(t)o(t)$ instead of the signal $o(t)$ for a window function $w(t)$. The triangle function is another popular window. Plot its Fourier transform. Compare it to the transform of the boxcar.

Exercise 2.3 The Welch window has the definition

$$w(t) = 1 - \frac{t^2}{p^2},$$

for $-p \leq t \leq p$, zero otherwise. Compute and plot its Fourier transform.

Exercise 2.4 The von-Hann window has the definition

$$w(t) = \frac{1}{2}\left(1 + \cos\frac{\pi t}{p}\right),$$

for $-p \leq t \leq p$, zero otherwise. Compute and plot its Fourier transform.

Exercise 2.5 The Blackman–Nuttall window has the definition

$$w(t) = 0.3635819 + 0.4891775\cos\frac{\pi t}{p} + 0.1365995\cos\frac{2\pi t}{p} + 0.0106411\cos\frac{3\pi t}{p},$$

for $-p \leq t \leq p$, zero otherwise. Compute and plot its Fourier transform.

Exercise 2.6 Consider the spring-mass-damper system in Fig. 2.3. Play with the damping constant c and observe the height and shape of the resonant peak.

Exercise 2.7 Derive the range of damping constants for which the spring-mass-damper system in Fig. 2.3 will exhibit resonance.

2.14 Lab Exercise

Exercise 2.8 Use a function generator to generate signals with different waveforms and an oscilloscope to visualize these signals. Familiarize yourself with the operation of these instruments. While exploring the modulation options of your signal generator observe both the resulting waveforms and the amplitudes spectrum of the waveform on your scope.

References

Franklin GF, Powell DJ, Emami-Naeini A (2010) Feedback control of dynamic systems, 6th edn. Pearson, Upper Saddle River

Franklin GF, Workman ML, Powell D (1997) Digital control of dynamic systems, 3rd edn. Addison-Wesley Longman Publishing Co., Inc., Boston

Gershenfeld N (1999) The nature of mathematical modeling. Cambridge University Press, Cambridge

Jänich K (2001) Analysis für Physiker und Ingenieure Funktionentheorie, Differentialgleichungen Spezielle Funktionen, Springer-Lehrbuch. Springer, Berlin

Kester W (2003) Mixed-signal and DSP design techniques, Analog Devices series. Analog Devices and Newnes, Amsterdam

Kreyszig E (2010) Advanced engineering mathematics. Wiley, New York

Lee EA, Varaiya P (2011) Structure and interpretation of signals and systems, 2nd edn. http://LeeVaraiya.org/

Middleton D (1996) An introduction to statistical communication theory: an IEEE press classic reissue. Wiley, New York

Oppenheim AV, Schafer R (1975) Digital signal processing. Prentice-Hall, Englewood Cliffs

Schoukens J, Renneboog J (1986) Modeling the noise influence on the Fourier coefficients after a discrete Fourier transform. IEEE Trans Instrum Meas 35(3):278–286

Zygmund A (2002a) Trigonometric series, vol I. Cambridge University Press, Cambridge

Zygmund A (2002b) Trigonometric series, vol II. Cambridge University Press, Cambridge

Springer

http://www.springer.com/978-3-319-10679-3

Engineering Embedded Systems
Physics, Programs, Circuits
Hintenaus, P.
2015, XXI, 345 p. 307 illus., 83 illus. in color., Hardcover
ISBN: 978-3-319-10679-3