Chapter 2

System Theory

Authors: Peter Fischer, Klaus Sembritzki, and Andreas Maier

In the digital age, any medical image needs to be transformed from continuous domain to discrete domain (i. e. 1's and 0's) in order to be represented in a computer. To do so, we have to understand what a **continuous** and a **discrete signal** is. Both of them are handled by **systems** which will also be introduced in this chapter. Another fundamental concept is the Fourier transform as it allows us to represent any time domain signal in frequency space. In particular, we will find that both representations – time domain and frequency domain – are equivalent and can be converted into each other. Having found this important relationship, we can then determine conditions which will guarantee that also conversion from continuous to discrete domain and vice versa is possible without loss of information. On the way, we will introduce several other important concepts that will also find repeated use later in this book.

A. Maier et al. (Eds.): Medical Imaging Systems, LNCS 11111, pp. 13–36, 2018.
https://doi.org/10.1007/978-3-319-96520-8_2

2.1 Signals and Systems

2.1.1 Signals

A signal is a function $f(t)$ that represents information. Often, the independent variable t is a physical dimension, like time or space. The output f of the signal is also called the dependent variable. Signals are everywhere in everyday life, although we are mostly not aware of them. A very prominent example is the speech signal, where the independent variable is time. The dependent variable is the electric signal that is created by measuring the changes of air pressure using a microphone. The description of the speech generation process enables to do efficient speech processing, e. g., radio transmission, speech coding, denoising, speech recognition, and many more. In general, many domains can be described using system theory, e. g., biology, society, economy. For our application, we are mainly interested in medical signals.

Both the dependent and the independent variable can be multidimensional. Multidimensional independent variables t are very common in images. In normal camera images, space is described using two spatial coordinates. However, medical images, e. g., CT volume scans, can also have three spatial dimensions. It is not necessary that all dimensions have the same meaning. Videos have two spatial coordinates and one time coordinate. In the medical domain, we can also find higher-dimensional examples like time-resolved 4-D MR and CT with three spatial dimensions and one time dimension. To represent multidimensional values, i. e., vectors, we use bold-face letters $\boldsymbol{t}$ or multiple scalar values, e. g., $\boldsymbol{t} = (x, y, z)^\mathsf{T}$. The medical field also contains examples of multidimensional dependent variables f. An example with many dimensions is the Electroencephalography (EEG). Electrodes are attached to the skull and measure electrical brain activity from multiple positions over time. To represent multidimensional dependent variables, we also use bold-face letters $\boldsymbol{f}$.

The signals described above are all in continuous domain, e. g., time and space change continuously. Also, the dependent variables vary continuously in principle, like light intensity and electrical voltage. However, some signals exist naturally in discrete domains w. r. t. the independent variable or the dependent variable. An example for a discrete signal in dependent and independent variable is the number of first semester students in medical engineering. The independent variable time is discrete in this case. The starting semesters are WS 2009, WS 2010, WS 2011, and so on. Other points in time are considered to be constant in this interval. The number of students is restricted to natural numbers. In general, it is also possible that only the dependent or the independent variable is discrete and the other one continuous. In addition to signals that are discrete by nature, other signals must be represented discretely for processing with a digital computer, which means that the independent variable must be discretized before processing with a com-

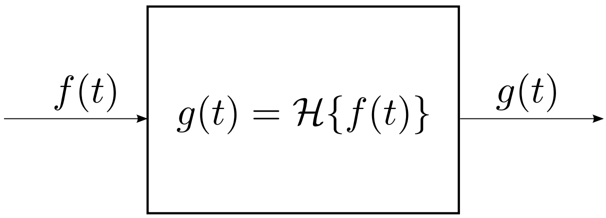


Figure 2.1: A system $\mathcal{H}\{.\}$ with the input signal $f(t)$ and the output signal $g(t)$.

puter. Furthermore, data storage in computers has limited precision, which means that the dependent variable must be discrete. Both are a direct consequence of the finite memory and processing speed of computers. This is the reason why discrete system theory is very important in practice.

Signals can be further categorized into deterministic and stochastic signals. For a deterministic signal, the whole waveform is known and can be written down as a function. In contrast, stochastic signals depend randomly on the independent variable, e. g., if the signal is corrupted by noise. Therefore, for practical applications, the stochastic properties of signals are very important. Nevertheless, deterministic signals are important to analyze the behavior of systems. A short introduction into stochastic signals and randomness will be given in Sec. 2.4.3.

This chapter is presents basic knowledge on how to represent, analyze, and process signals. The correct processing of signals requires some math and theory. A more in-depth introduction into the concepts presented here can be found in [3]. The application to medical data is treated in [2].

2.1.2 Systems

Signals are processed in processes or devices, which are abstracted as **systems**. This includes not only technical devices, but natural processes like attenuation and reverberation of speech in transmission through air as well. Systems have signals as input and as output. Inside the system, the properties of the signal are changed or signals are related to each other. We describe the processing of a signal using a system with the operator $\mathcal{H}\{\cdot\}$ that is applied to the function f. A graphical representation of a system is shown in Fig. 2.1.

An important subtype is the **linear shift-invariant system**. Linear shift-invariant systems are characterized by the two important properties of linearity and shift-invariance (cf. Geek Box 2.1 and 2.2).

Another property important for the practical realization of linear shift-invariant systems is causality. A causal system does not react to the input

Geek Box 2.1: Linear Systems

The linearity property of a system means that linear combinations of inputs can be represented as the same linear combination of the processed inputs

$$\mathcal{H}\{af(t)\} = a\mathcal{H}\{f(t)\} \tag{2.1}$$
$$\mathcal{H}\{f(t) + g(t)\} = \mathcal{H}\{f(t)\} + \mathcal{H}\{g(t)\}, \tag{2.2}$$

with constant a and arbitrary signals f and g. The linearity property greatly simplifies the mathematical and practical treatment, as the behavior of the system can be studied on basic signals. The behavior on more complex signals can be inferred directly if they can be represented as a superposition of the basic signals.

Geek Box 2.2: Shift-Invariant Systems

Shift-invariance denotes the characteristic of a system that its response is independent of shifts of the independent variable of the signal. Mathematically, this is described as

$$g_1(t) = \mathcal{H}\{f(t)\} \tag{2.3}$$
$$g_2(t) = \mathcal{H}\{f(t-\tau)\} \tag{2.4}$$
$$g_1(t-\tau) = g_2(t), \tag{2.5}$$

for the shift τ. This means that shifting the signal by τ followed by processing with the system is identical to processing the signal with the system followed by a shift with τ.

before the input actually arrives in the system. This is especially important for signals with time as the independent parameter. However, non-causal systems do not pose a problem for the independent parameter space, e. g., image filters that use information from the left and right of a pixel. Geek Box 2.3 presents examples for the combination of different system properties.

Linear shift-invariant systems are important in practice and have convenient properties and a rich theory. For linear shift-invariant systems, the abstract operator $\mathcal{H}\{\cdot\}$ can be described completely using the impulse response $h(t)$ (cf. Sec. 2.2.2) or transfer function $H(\xi)$ (cf. Sec. 2.3.2). The impulse response is combined with the signal by the operation of convolution. This is sufficient to describe all linear shift-invariant systems.

Geek Box 2.3: System Examples

Here are some examples of different systems analyzed w. r. t. linearity, shift-invariance, and causality. $f(t)$ represents the input and $g(t)$ the output signal.

- $g(t) = 10f(t)$: linear, shift-invariant, causal
- $g(t) = \sin(f(t))$: non-linear, shift-invariant, causal
- $g(t) = 3f(t+2)$: linear, shift-invariant, non-causal
- $g(t) = f(t) - 2f(t-1)$: linear, shift-invariant, causal
- $g(t) = f(t) \cdot \mathrm{e}^{(-0.5t)}$: linear, not shift-invariant, causal

2.2 Convolution and Correlation

This section describes the combination of signals in linear-shift-invariant systems, i. e., convolution or correlation. Before discussing signal processing in detail, we will first start by revisiting important mathematical concepts that will be needed in the following chapters.

2.2.1 Complex Numbers

Complex numbers are an extension to real numbers. They are defined as $z = a + bi$. a is called the real part of z and b the imaginary part. Both act as coordinates in a 2-D space. i is the imaginary unit that spans the second dimension of this space. The special meaning of i is that $i^2 = -1$. This makes complex numbers important for many areas in mathematics, but also in many applied fields like physics and electrical engineering. To extract the coordinates of the complex number, we use the following definitions

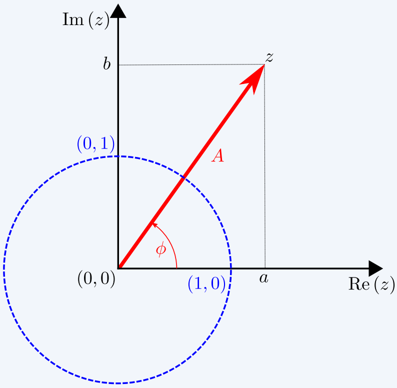
$$a = \mathrm{Re}\,(z) \tag{2.6}$$

$$b = \mathrm{Im}\,(z)\,. \tag{2.7}$$

We can directly write $z = \mathrm{Re}\,(z) + \mathrm{Im}\,(z)\,i$. Another important definition is the complex conjugate $\bar{z}$, which is the same number as z except with the opposite sign for the imaginary part $\bar{z} = a - bi$.

Real numbers are the subset of the complex numbers for which $b = 0$, i. e., no imaginary part. Geometrically, this means that real numbers are defined on a one-dimensional axis, whereas the complex numbers are defined on a 2-D plane. The geometric interpretation of complex numbers is also helpful to see the equivalence of the Cartesian coordinate notation $z = a + bi$ and the polar coordinate notation $z = A\,(\cos\phi + i \sin\phi)$ of complex numbers. The

Geek Box 2.4: Complex Numbers and Geometric Interpretation



If a point on the 2-D plane is seen as a position vector, A is the length of the vector and ϕ the angle relative to the real axis. The two notations can be converted to each other using the following formulas:

$$A = \sqrt{a^2 + b^2}$$

$$\phi = \begin{cases} \arctan \frac{b}{a}, & \text{if } a > 0 \\ \arctan \frac{b}{a} + \pi, & \text{if } a < 0 \text{ and } b \geq 0 \\ \arctan \frac{b}{a} - \pi, & \text{if } a < 0 \text{ and } b < 0 \\ \frac{\pi}{2}, & \text{if } a = 0 \text{ and } b > 0 \\ -\frac{\pi}{2}, & \text{if } a = 0 \text{ and } b < 0 \\ \text{undefined}, & \text{if } a = 0 \text{ and } b = 0 \end{cases}$$

$$a = A \cos \phi$$

$$b = A \sin \phi$$

polar coordinates consists of magnitude A and angle ϕ (cf. Geek Box 2.4). For system theory, an important property of complex numbers is Euler's formula

$$\exp(i\phi) = e^{i\phi} = \cos(\phi) + i \sin(\phi). \tag{2.8}$$

Using this relation, a complex sum of sine and cosine can be expressed conveniently using a single exponential function. This leads directly to the exponential notation of complex numbers $z = Ae^{i\phi}$. We will use the complex numbers and different notations in Sec. 2.3.

Description	Equation
Linearity	$g(t) * (a \cdot f(t) + b \cdot h(t)) = a((g * f)(t)) + b((g * h)(t))$
Shift-invariance	$g(t) * f(t - \tau) = (g * f)(t - \tau)$
Commutativity	$g(t) * f(t) = f(t) * g(t)$
Associativity	$g(t) * ((f * h)(t)) = ((f * g)(t)) * h(t)$
Distributivity	$f(t) * (g(t) + h(t)) = (f * g)(t) + (f * h)(t)$

Table 2.1: Some mathematical properties of convolution. a, b are constants.

2.2.2 Convolution

As mentioned above, convolution is the operation that is necessary to describe the processing of any signal with a linear shift-invariant system. Convolution in the continuous case is defined as

$$g(t) = (h * f)(t) = \int_{-\infty}^{\infty} h(\tau) f(t - \tau) \, \mathrm{d}\tau. \tag{2.9}$$

In order for the convolution to be well-defined, some requirements for the functions h and f must be fulfilled. For the infinite integral to exist, h and f must decay fast enough towards infinity. This is the case if one of the functions has compact support, i. e., it is 0 everywhere except for a limited region. As an example, the convolution of a square input function $f(t)$ with an Gaussian function $h(t)$ is investigated in Geek Box 2.5. Further mathematical properties of convolution are listed in Table 2.1.

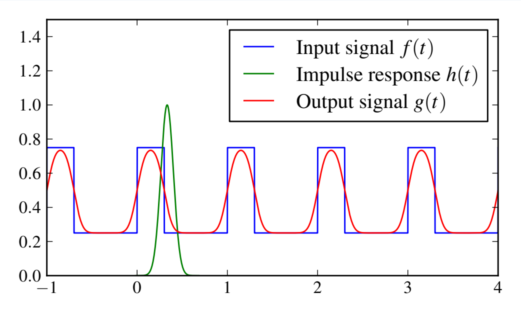
A common basic signal is the Dirac function which is also called delta function or impulse function. It is a infinitely short, infinitely high impulse.

$$\delta(t) = \begin{cases} \infty, & \text{if } t = 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.10}$$

It is impossible to describe the Dirac function using classical functions. It requires the use of generalized functions or distributions, which is out of the scope of this introduction. The Dirac function is usually represented graphically as an arrow of length 1, see Fig. 2.2.

Sequences of Dirac pulses are useful to select only certain points of a function like a sifter (cf. Figure 2.3). The sifting property of the Dirac function is given by integrating the product of a function and a time-delayed Dirac function

Geek Box 2.5: Convolution Example



For the definition of the square function, the Heaviside step function is useful to shorten the notation

$$H(t) = \begin{cases} 0, & \text{if } t < 0 \\ 1, & \text{otherwise} \end{cases}.$$

Then, the square function and the Gaussian are defined as

$$f(t) = k_1 + k_2 \sum_{n=-\infty}^{\infty} H(t - nT) - H(t - nT - k_3)$$
$$h(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2},$$

with the offset k_1, the amplitude k_2, the duty-cycle k_3, and the period T of the square function and the standard deviation σ of the Gaussian. The convolution with a Gaussian results in a smoothing of the edges of the square function.

$$\int_{-\infty}^{\infty} f(t)\delta(t - T)\,\mathrm{d}t = f(T).$$

With the sifting property, the element at $t = T$ can be selected from the function, which is equivalent to sampling the function at that time point.

The sift property is useful for convolution of an arbitrary function and the Dirac function.

$$f(t) * \delta(t - T) = \int_{-\infty}^{\infty} f(\tau)\delta(t - T - \tau)\,\mathrm{d}\tau = f(t - T) \tag{2.11}$$

Consequently, the Dirac function is the identity element of convolution.

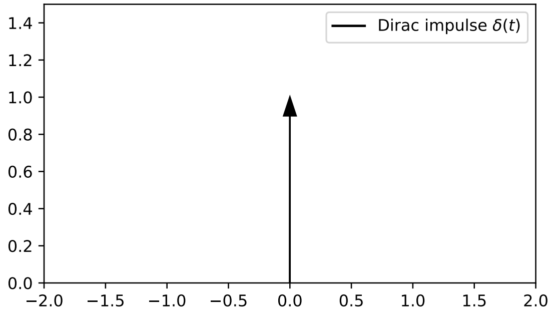


Figure 2.2: Graphical representation of the Dirac function $\delta(t)$. The arrow symbolizes infinity.

Figure 2.3: Laboratory sifters are used to remove undesired parts from discrete signals. Sequences of Dirac pulses can be applied in a similar way. Image courtesy of BMK Wikimedia.

The response of a system to a Dirac function on the input is called the impulse response of the system $h(t) = \mathcal{H}\{\delta(t)\}$. Using the superposition principle, every other signal can be represented as a linear combination of infinitely many Dirac functions. Therefore, the output of a system to any input signal is computed by convolution of the input signal $f(t)$ with the impulse response $h(t)$.

$$g(t) = f(t) * h(t) \tag{2.12}$$

For medical applications, an important example of a linear shift-invariant system is an imaging system. The output of an imaging system is often modeled as a linear shift-invariant system. The impulse response of an imaging system is called point spread function. It describes how a single point, i. e., a Dirac impulse, is spread on the sensor plane by the specific imaging system. The point spread function is a description of the behavior of the system.

2.2.3 Correlation

Another basic operation to combine a signal and a system is correlation

$$g(t) = (h \star f)(t) = \int_{-\infty}^{\infty} h^*(\tau) f(t+\tau)\,\mathrm{d}\tau, \tag{2.13}$$

where h^* is the complex conjugate of h. The main difference to convolution is that the input signal f is not mirrored before combination with h, i. e., $f(t+\tau)$ instead of $f(t-\tau)$. Correlation is a way to measure the similarity of two signals.

An application of correlation is the matched filter. The matched filter is specifically designed to have a high response for a specific deterministic signal or waveform $f(t)$. It is **matched** to that signal. The matched filter is directly computed by correlation with the desired signal. Alternatively, convolution with an impulse response of the mirrored, complex conjugate of the desired deterministic signal $h(t) = f^*(-t)$ can be used.

Technical uses for correlation can be found in signal transmission and signal detection. For a medical example, the heartbeats of a person can be detected in an Electrocardiogram (ECG) using correlation with a template QRS complex (QRS complex denotes the combination of three of the graphical deflections seen on an ECG). In image processing, a certain deterministic signal is searched for across the whole image. In this case, the deterministic signal is often called template and the process of searching is called template matching. This can be used for the detection of specific structures and tracking of structures over time. Geek Box 2.6 puts the correlation in signal processing in relation to the statistical correlation coefficient.

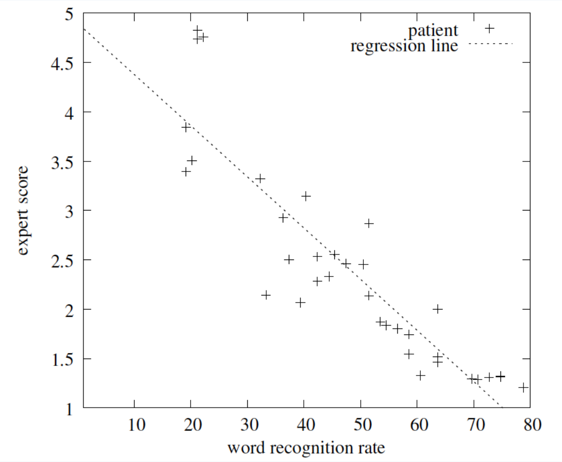
2.3 Fourier Transform

Up to this point, all operations and mathematical definitions were performed in continuous domain. Also, we have not discussed the relation between discrete and continuous representations which are important to understand the concept of sampling. In the following, we will introduce the Fourier transform and related concepts which will allow us to deal with exactly such problems.

2.3.1 Types of Fourier Transforms

A cosine wave f of time t with amplitude A, frequency ξ, and phase shift φ can be described by the following three equivalent parametrizations.

Geek Box 2.6: Relation to the Statistical Correlation Coefficient



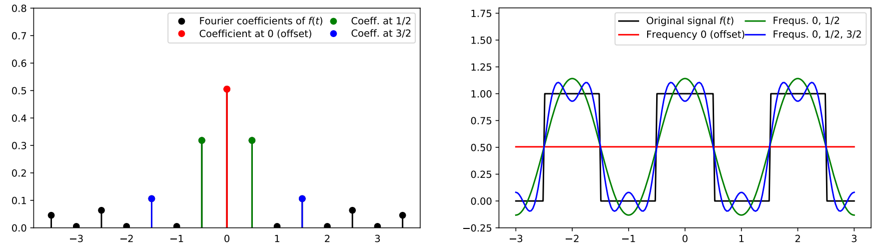
In statistics, the so-called Pearson correlation coefficient r [5] is a measure of agreement between two sets of observations $\boldsymbol{x}$ and $\boldsymbol{y}$. Coefficient r is defined in the interval $[-1, 1]$ and if $|r| = 1$, a perfect linear relationship between the two variables is present. It is computed in the following way:

$$r(\boldsymbol{x}, \boldsymbol{y}) = \frac{\sum_n (x_n - \bar{x})(y_n - \bar{y})}{\sigma_x \sigma_y}$$

Here, we use $\bar{x}$, $\bar{y}$, σ_x, and σ_y to denote the respective mean values and standard deviations. If we assume the standard deviations to be equal to 1 and the means equal to 0, we arrive at the following equation:

$$r(\boldsymbol{x}, \boldsymbol{y}) = \sum_n x_n \cdot y_n$$

This is identical to the discrete version of correlation for real inputs for $t = 0$. Also note that this can be considered simply as an inner product $\boldsymbol{x}^\top \boldsymbol{y}$.
The image at the top of the page shows a scatter plot between two variables *word recognition rate* and *expert rater*. Each point (x_n, y_n) denotes one patient for whom both of the two variables were measured. The closer the two are to the dotted line, the better their agreement. Here, their dependency is negative as if one variable is high, the other is low and vice versa. $r \approx -0.9$ in this example. Please refer to [4] for more details.



(a) Fourier coefficients, weights of trigonometric functions approximating the signal $f(t)$

(b) Periodic signal and approximations using different numbers of Fourier coefficients

Figure 2.4: Approximation of a periodic signal using a weighted sum of trigonometric functions

$$
\begin{aligned}
f(t) &= A \cdot \cos(2\pi\xi t + \varphi) && A, \varphi \in \mathbb{R} \\
&= a \cdot \cos(2\pi\xi t) + b \cdot \sin(2\pi\xi t) && a, b \in \mathbb{R} \\
&= c \cdot \mathrm{e}^{2\pi i\xi t} + \bar{c} \cdot \mathrm{e}^{-2\pi i\xi t} && c \in \mathbb{C}
\end{aligned}
$$

In Geek Box 2.7, we show how the parameters a, b, and c are related to A and φ.

A **Fourier series** (cf. Geek Box 2.8) is used to represent a continuous signal using only discrete frequencies. As such a Fourier series is able to approximate any signal as a superposition of sine and cosine waves. Fig. 2.4(b) shows a rectangular signal of time. The absolute values of its Fourier coefficients are depicted in Fig. 2.4(a). As can be seen in Fig. 2.4(a), the Fourier coefficients decrease as the frequency increases. It is therefore possible to approximate the signal by setting the coefficients to 0 for all high frequencies. Fig. 2.4(b) includes the approximations for three different choices of sets of frequencies.

The Fourier series, which works on periodic signals, can be extended to aperiodic signals by increasing the period length to infinity. The resulting transform is called **continuous Fourier transform** (or simply Fourier transform, cf. Geek Box 2.9). Fig. 2.5(b) shows the Fourier transform of a rectangular function, which is identical to the Fourier coefficients at the respective frequencies up to scaling (see Fig. 2.5(a)).

The counter part to the Fourier series for cases in which time domain is discrete and the frequency domain is continuous is called the **discrete time Fourier transform** (cf. Geek Box 2.10). It forms a step towards the **discrete Fourier transform** (cf. Geek Box 2.11) which allows us to perform all previous operations also in a digital signal processing system. In discrete space, we can interpret the Fourier transform simply as a matrix multiplication with a complex matrix $\boldsymbol{F}$

Name	Function	Fourier transform
Rectangular	$\mathrm{rect}(at) = \begin{cases} 0 & \text{if } \lvert at\rvert > \frac{1}{2} \\ \frac{1}{2} & \text{if } \lvert at\rvert = \frac{1}{2} \\ 1 & \text{if } \lvert at\rvert < \frac{1}{2} \end{cases}$	$\mathcal{F}\left[\mathrm{rect}(t)\right](\xi) = \frac{1}{\lvert a\rvert}\,\mathrm{sinc}(\frac{\xi}{a})$
Triangular	$\mathrm{tri}(t) = \begin{cases} 1 - \lvert t\rvert & \text{if } \lvert t\rvert < 1 \\ 0 & \text{if } \lvert t\rvert \leq 1 \end{cases}$	$\mathcal{F}\left[\mathrm{tri}(t)\right](\xi) = \mathrm{sinc}^2(\xi)$
Gaussian	$\mathrm{gauss}(t) = \mathrm{e}^{-at^2}$	$\mathcal{F}\left[\mathrm{gauss}(t)\right](\xi) = \sqrt{\frac{\pi}{a}}\mathrm{e}^{-\pi^2\xi^2/a}$

Table 2.2: Fourier transforms of popular functions. Here we use the definition $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$. Note that a convolution of two rectangular functions yields a triangular function as $\mathcal{F}\left[\mathrm{rect}(t) * \mathrm{rect}(t)\right] = \mathrm{sinc}^2(\xi)$.

$$\boldsymbol{k} = \boldsymbol{F}\boldsymbol{n} \tag{2.14}$$

where the signal $\boldsymbol{n}$ and the discrete spectrum $\boldsymbol{k}$ are vectors of complex values. The inverse operation is then readily found as

$$\boldsymbol{n} = \boldsymbol{F}^H\boldsymbol{k} \tag{2.15}$$

where $\boldsymbol{F}^H$ is the Hermitian, i. e., transposed and element-wise conjugated, of $\boldsymbol{F}$. Geek Box 2.12 shows some more details on how to find these relations. Fig. 2.5 shows all types of Fourier transforms introduced in this section in comparison. Tab. 2.2 shows the Fourier transforms of popular functions.

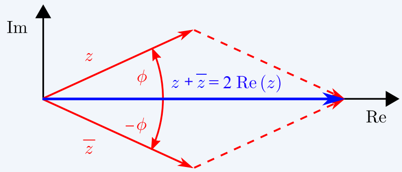
In computer programs, discrete Fourier transforms are implemented very efficiently using fast Fourier transform (FFT). This approach reduces the number of computations from the order of N^2 to the order of $N \log N$, if N is the length of the signal. In the next section, we will see why convolution and correlation also benefit from this efficiency.

2.3.2 Convolution Theorem & Properties

The convolution of two functions f and g is defined as in Sec. 2.2.2, and $\cdot$ denotes point-wise multiplication. The convolution theorem states that a convolution of two signals in space is identical to a point-wise multiplication of their spectra (see Equation 2.24). The opposite also holds true (see Equation 2.25).

$$\mathcal{F}\{f * g\} = F \cdot G \tag{2.24}$$
$$\mathcal{F}\{f \cdot g\} = F * G \tag{2.25}$$

Geek Box 2.7: Equivalent Cosine Representations



Oscillations of the same frequency can be represented in several equivalent ways. In the following, we make use of the complex numbers introduced in Sec. 2.2.1 and the correspondence between a sum of complex exponentials and the real part $z+\overline{z}=2\,\mathrm{Re}(z)$ to convert the different representations into the same expression.

Amplitude and phase shift, where we define $c=\frac{1}{2}A\mathrm{e}^{i\varphi}$:

$$\begin{aligned} f(t) &= A\cdot\cos(2\pi\xi t+\varphi) = \mathrm{Re}\left(A\cdot \mathrm{e}^{2\pi i\xi t+i\varphi}\right) \\ &= \mathrm{Re}\left(A\cdot \mathrm{e}^{i\varphi}\cdot \mathrm{e}^{2\pi i\xi t}\right) = \underline{\underline{\mathrm{Re}\left(2c\cdot \mathrm{e}^{2\pi i\xi t}\right)}}. \end{aligned}$$

Sum of cosine and sine functions, where we define $c=\frac{1}{2}(a-ib)$:

$$\begin{aligned} f(t) &= a\cdot\cos(2\pi\xi t)+b\cdot\sin(2\pi\xi t) \\ &= a\cdot\cos(2\pi\xi t)+b\cdot\cos(2\pi\xi t-\pi/2) \\ &= \mathrm{Re}\left(a\cdot \mathrm{e}^{2\pi i\xi t}\right)+\mathrm{Re}\left(b\cdot \mathrm{e}^{2\pi i\xi t-\pi/2}\right) \\ &= \mathrm{Re}\left(a\cdot \mathrm{e}^{2\pi i\xi t}\right)+\mathrm{Re}\left(b\cdot \mathrm{e}^{2\pi i\xi t}\cdot \mathrm{e}^{-i\pi/2}\right) \\ &= \mathrm{Re}\left(a\cdot \mathrm{e}^{2\pi i\xi t}\right)+\mathrm{Re}\left(-ib\cdot \mathrm{e}^{2\pi i\xi t}\right) \\ &= \mathrm{Re}\left((a-ib)\cdot \mathrm{e}^{2\pi i\xi t}\right) = \underline{\underline{\mathrm{Re}\left(2c\cdot \mathrm{e}^{2\pi i\xi t}\right)}}. \end{aligned}$$

Sum of complex exponentials:

$$\begin{aligned} f(t) &= c\cdot \mathrm{e}^{2\pi i\xi t}+\bar{c}\cdot \mathrm{e}^{-2\pi i\xi t} \\ &= \mathrm{Re}\left(c\cdot \mathrm{e}^{2\pi i\xi t}\right)+i\,\mathrm{Im}\left(c\cdot \mathrm{e}^{2\pi i\xi t}\right)+\mathrm{Re}\left(c\cdot \mathrm{e}^{2\pi i\xi t}\right)-i\,\mathrm{Im}\left(c\cdot \mathrm{e}^{2\pi i\xi t}\right) \\ &= \underline{\underline{\mathrm{Re}\left(2c\cdot \mathrm{e}^{2\pi i\xi t}\right)}}. \end{aligned}$$

Geek Box 2.8: Fourier Series

The Fourier series (Equation 2.17) represents a periodic signal of period T by an infinite weighted sum of shifted cosine functions of different frequencies. The Fourier coefficients c are calculated using Equation 2.16.

$$c[k] = \frac{1}{T} \int_{d}^{d+T} f(t)\ \mathrm{e}^{-2\pi i t k/T}\, \mathrm{d}t \qquad k \in \mathbb{Z} \tag{2.16}$$

$$f(t) = \sum_{k=-\infty}^{\infty} c[k]\ \mathrm{e}^{2\pi i t k/T} \qquad t \in \mathbb{R} \tag{2.17}$$

The coefficients $c[k]$ and $c[-k]$ together form a shifted cosine wave with frequency $\xi = \frac{|k|}{T}$ (see Geek Box 2.7). It follows that $c[-k] = \overline{c[k]}$:

$$\begin{aligned} c[k]\ \mathrm{e}^{2\pi i t k/T} + c[-k]\ \mathrm{e}^{-2\pi i t k/T} &= c[k]\ \mathrm{e}^{2\pi i t k/T} + \overline{c[k]}\ \mathrm{e}^{-2\pi i t k/T} \\ c[-k]\ \mathrm{e}^{-2\pi i t k/T} &= \overline{c[k]}\ \mathrm{e}^{-2\pi i t k/T} \\ \Rightarrow c[-k] &= \overline{c[k]} \end{aligned}$$

Geek Box 2.9: Continuous Fourier Transform

Given a time-dependent signal f, its Fourier transform F at frequency ξ is defined by Eq. (2.18). The inverse Fourier transform is defined by Eq. (2.19).

$$F(\xi) = \int_{-\infty}^{\infty} f(t)\ \mathrm{e}^{-2\pi i t \xi}\, \mathrm{d}t \qquad \xi \in \mathbb{R} \tag{2.18}$$

$$f(t) = \int_{-\infty}^{\infty} F(\xi)\ \mathrm{e}^{2\pi i t \xi}\, \mathrm{d}\xi \qquad t \in \mathbb{R} \tag{2.19}$$

In general, $f(t)$ can be a complex signal. We will, however, only consider the case where $f(t)$ is real-valued. The continuous Fourier transform is symbolized by the operator $\mathcal{F}$.

Geek Box 2.10: Discrete-time Fourier Transform

The spectrum (i. e., continuous Fourier transform) of a band-limited signal that is sampled equidistantly and sufficiently dense with distance T can be calculated using the discrete-time Fourier transform (DTFT) defined by Equation 2.20. The inverse transform is given by Equation 2.21. For details about the required sampling distance see Sec. 2.4.2.

$$F_{\frac{1}{T}}(\xi) = \sum_{n=-\infty}^{\infty} f[n]\ \mathrm{e}^{-2\pi i \xi n T} \qquad \xi \in \mathbb{R} \tag{2.20}$$

$$f[n] = T \int_{d}^{d+\frac{1}{T}} F_{\frac{1}{T}}(\xi)\ \mathrm{e}^{2\pi i \xi n T}\,\mathrm{d}\xi \qquad n \in \mathbb{Z} \tag{2.21}$$

Fig. 2.5(c) shows the DTFT of a band-limited function and the Fourier transform. The DTFT is identical to the Fourier transform up to scaling except that it is periodic with period $1/T$.

Geek Box 2.11: Discrete Fourier Transform

The spectrum of a periodic and band-limited signal can be calculated with the discrete Fourier transform (DFT) as defined by Equation 2.22. The signal can be reconstructed with the inverse DFT as defined by Equation 2.23.

$$F[k] = \sum_{n=0}^{N-1} f[n]\ \mathrm{e}^{-2\pi i n k/N} \qquad k \in \mathbb{Z} \tag{2.22}$$

$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} F[k]\ \mathrm{e}^{2\pi i n k/N} \qquad n \in \mathbb{Z} \tag{2.23}$$

Fig. 2.5(d) shows the DFT and the Fourier series of a band-limited signal. The DFT is identical to the Fourier series up to scaling except that it is periodic with period $1/N$.

Geek Box 2.12: Discrete Fourier Transform as Matrix

A discrete Fourier transform can be rewritten as a complex matrix product. To demonstrate this, we start with the definition of the discrete Fourier transform:

$$\begin{aligned} F[k] &= \sum_{n=0}^{N-1} f[n]\ \mathrm{e}^{-2\pi ink/N} \\ &= \sum_{n=0}^{N-1} \mathrm{e}^{-2\pi ink/N}\ f[n] \end{aligned}$$

Now, we replace the summation with an inner product of two vectors $\boldsymbol{\xi}_k$ and $\boldsymbol{n}$ (cf. Geek Box 2.6):

$$F[k] = (\mathrm{e}^0, \mathrm{e}^{-2\pi nk/N}, \ldots, \mathrm{e}^{-2\pi i(N-1)k/N}) \begin{pmatrix} f[0] \\ f[1] \\ f[2] \\ \vdots \\ f[N-1] \end{pmatrix} = \boldsymbol{\xi}_k^\top \boldsymbol{n}$$

We see that $\boldsymbol{\xi}_k$ is a discretely sampled wave at frequency k. This equation can now be interpreted as the k-th row of a matrix vector product. Thus, we can rewrite the entire discrete Fourier transform of all K frequencies to

$$\boldsymbol{k} = \begin{pmatrix} F[0] \\ F[1] \\ \vdots \\ F[K-1] \end{pmatrix} = \begin{pmatrix} \boldsymbol{\xi}_0^\top \\ \boldsymbol{\xi}_1^\top \\ \vdots \\ \boldsymbol{\xi}_{K-1}^\top \end{pmatrix} \boldsymbol{n} = \boldsymbol{F}\boldsymbol{n}$$

As such, each row of the above matrix multiplication computes a correlation between a wave of frequency k for all K frequencies under consideration. Furthermore the relation $\boldsymbol{F}^H = \boldsymbol{F}^{-1}$ holds if $\boldsymbol{F}^H$ is scaled with $\frac{1}{N}$. Hence, $\boldsymbol{F}$ forms an orthonormal basis. If we continue this line of thought, we can also interpret a Fourier transform as a basis rotation. In our case, we do not rotate by a certain angle, but we project our time-dependent signal into a frequency resolved time-independent space.

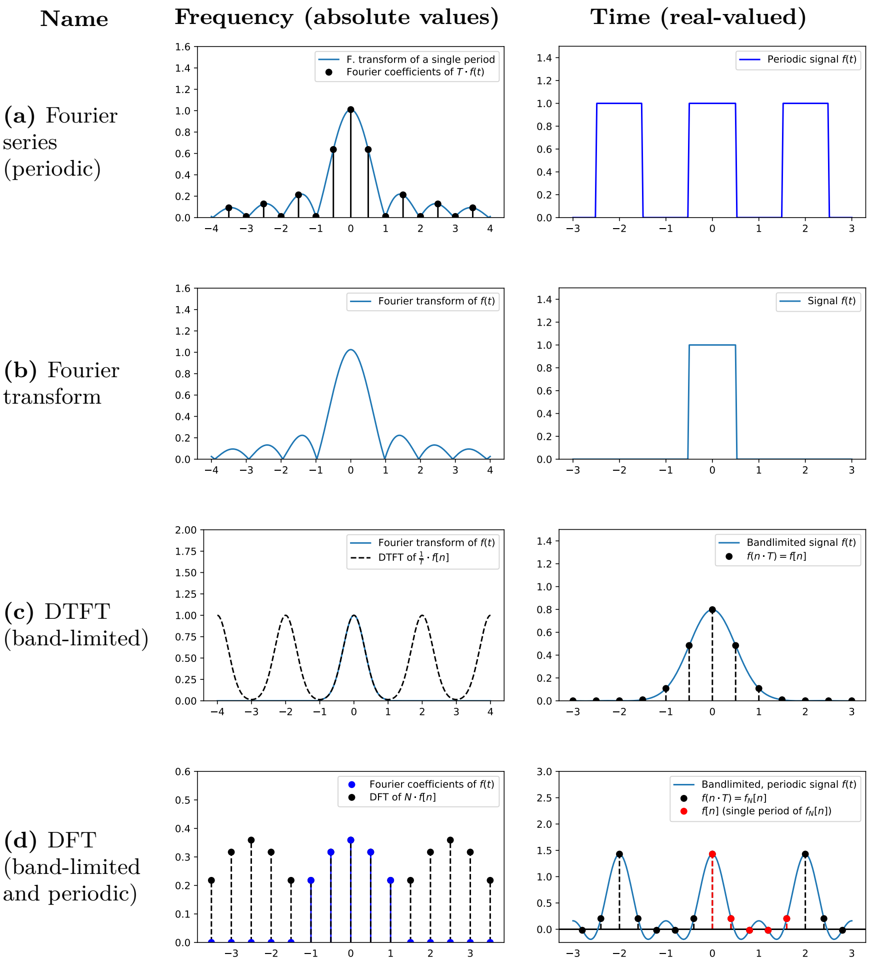


Figure 2.5: Different types of Fourier transforms.

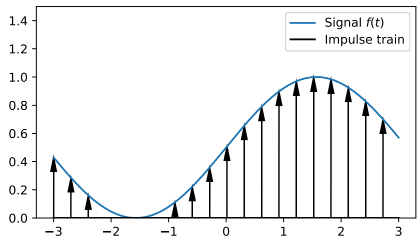
A similar theorem exists for the DFT. Let $\boldsymbol{C_h}$ denote the matrix that performs the convolution with discrete impulse response $\boldsymbol{h}$, and $\boldsymbol{f}$ be a discrete input signal. Then system output $\boldsymbol{g}$ is obtained as

$$\boldsymbol{g} = \boldsymbol{h} * \boldsymbol{f} = \boldsymbol{C}_h \boldsymbol{f} = \boldsymbol{F}^H \boldsymbol{H} \boldsymbol{F} \boldsymbol{f}.$$

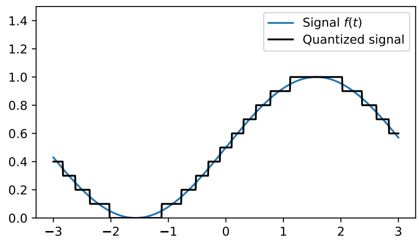
where $\boldsymbol{H}$ is a diagonal matrix that contains the Fourier transformed coefficients of $\boldsymbol{h}$. Note that $\boldsymbol{F}$ and $\boldsymbol{F}^H$ can be implemented efficiently by means of FFT. In addition to the convolution theorem, the Fourier transform has other notable properties. Some of those properties are listed in Table 2.3.

Description	Time	Frequency
Linearity	$a \cdot f(t) + b \cdot g(t)$	$a \cdot F(\xi) + b \cdot G(\xi)$
Shift	$f(t-a)$	$e^{-2\pi i a \xi} F(\xi)$
Scaling	$f(at)$	$\frac{1}{\lvert a \rvert} F\left(\frac{\xi}{a}\right)$
Derivative	$\frac{d^n f(t)}{dt^n}$	$(2\pi i \xi)^n F(\xi)$
Convolution theorem (see Sec. 2.3.2)	$(f * g)(t)$	$F(\xi) \cdot G(\xi)$
Dual of the convolution theorem	$f(t) \cdot g(t)$	$(F * G)(\xi)$

Table 2.3: Effects of modifications of a signal in time on the Fourier transform.



(a) Sampling as convolution with an impulse train.



(b) Quantization of a continuous signal to discrete values.

Figure 2.6: Discrete system theory

2.4 Discrete System Theory

2.4.1 Motivation

As already indicated in the introduction, discrete signals and systems are very important in practice. All signals can only be stored and processed at fixed discrete time instances in a digital computer. The process of transforming a continuous time signal to a discrete time signal is called **sampling**. In the simplest and most common case, the continuous signal is sampled at regular intervals, which is called uniform sampling. The current value of the continuous signal is stored exactly at the time instance where the discrete time signal is defined. This can be modeled by a convolution with an impulse train, see Fig. 2.6(a). At first glance, it looks like a lot of information is discarded in the process of sampling. However, under certain requirements, the continuous time signal can be reconstructed exactly. Further details are given in Sec. 2.4.2.

As we have already seen with the discrete Fourier transform, most methods introduced in this Chapter can be equally applied to discrete signals. We denote discrete signals using brackets [] instead of parentheses (), as we already did in the Geek Boxes. Integrals must be replaced by infinite sums, for example for the discrete convolution

$$g[n] = (h * f)[n] = \sum_{k=-\infty}^{\infty} h[k]f[n-k]. \tag{2.26}$$

In the discrete case, the Dirac function takes on a simple form.

$$\delta[n] = \begin{cases} 1, & \text{if } n = 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.27}$$

Note that in contrast to the continuous Dirac function, it is possible to exactly represent and implement the discrete Dirac function.

In addition to the discrete independent variable, the dependent variable can also be discrete. This means that the signal value $f(t)$ or $f[n]$ can only take values of certain levels. Apart from naturally discrete signals, all signals must be converted to a fixed discrete value for representation and processing in digital computers. For example, image intensities are often represented in the computer using 8 bit, i. e., 256 different intensities, or 12 bit which corresponds to 4096 different levels. The process of transforming a continuous-valued signal to a discrete-valued signal is called **quantization**. In most cases, a uniform quantization is sufficient, which means that the discrete levels have equal distance from each other. The continuous-valued signal is rounded to the nearest discrete level available, see Fig. 2.6(b). The error arising during this process is called quantization noise. Some more details on noise and noise models are given in Sec. 2.4.3.

2.4.2 Sampling Theorem

The Nyquist-Shannon sampling theorem (or just sampling theorem) states that a band-limited signal, i. e., a signal where all frequencies above ξ_B and below $-\xi_B$ are zero, can be fully reconstructed using samples $1/(2\xi_B)$ apart. If we consider a sine wave of frequency ξ_B, we have to sample it at least with a frequency of $2\,\xi_B$, i. e. twice per wavelength.

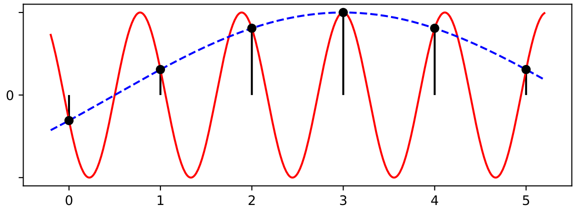


Figure 2.7: Sampling a sine signal with a frequency below $2\,\xi_B$ will cause aliasing. The reconstructed sine wave shown with blue dashes does not match the original frequency shown in red.

Formally, the theorem can be derived using the periodicity of the DTFT (see Fig. 2.5(c)). The DTFT spectrum is a periodic summation of the original spectrum, and the periodic spectra do not overlap as long as the sampling theorem is fulfilled. It is therefore possible to obtain the original spectrum by setting the DTFT spectrum to zero for frequencies larger than B. The signal can then be reconstructed by applying the inverse Fourier transform. We refer to [3] for a more detailed description of this topic.

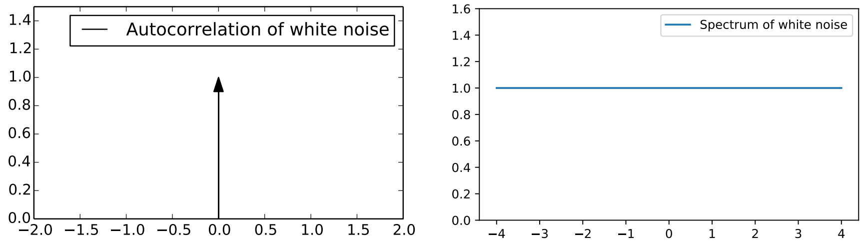
So far, we have not discussed how the actual sampling frequency $2\,\xi_B$ is determined. Luckily such a band limitation can be found for most applications. For example, even the most sensitive ears cannot perceive frequencies above 22 kHz. As a result, the sampling frequency of the compact disc (CD) was determined at 44.1 kHz. For the eye, typically 300 dots per inch in printing or 300 pixels per inch for displays are considered as sufficient to prevent any visible distortions. In videos and films, a frame rate of 50 Hz is often used to diminish flicker. High fidelity devices may support up to 100 Hz.

If the sampling theorem is not respected, *aliasing* occurs. Frequencies above the Nyquist frequency are wrapped around due to the periodicity and appear as lower frequencies. Then, these high frequencies are indistinguishable from the true low frequencies. Fig. 2.7 demonstrates this effect visually.

2.4.3 Noise

In many cases, acquired measurements or images are corrupted by some unwanted signal components. Common noise sources are quantization and thermal noise. Additional noise sources occur in the field of medical imaging, due to the related image acquisition techniques.

We can often find a simple model of the noise corrupting the image. The model does not represent the physical noise causes, but it approximately describes the errors that occur in the final signal. An additive noise model is commonly denoted as



(a) Autocorrelation **(b)** Spectrum

Figure 2.8: Example of a white noise function

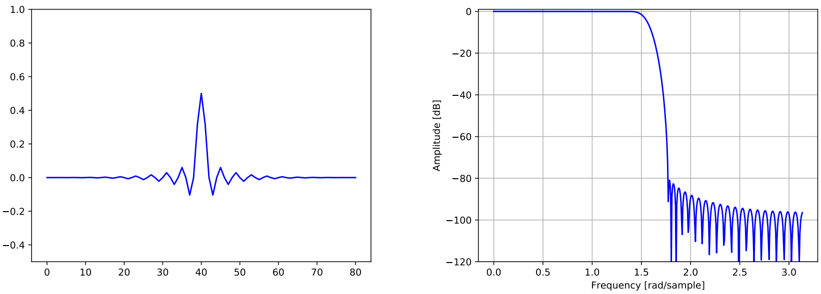
$$f(t) = s(t) + n(t) \tag{2.28}$$

where $s(t)$ is the underlying desired signal. We observe the signal $f(t)$, which is corrupted by the noise $n(t)$. For the statistics of the noise, we can use various models e. g., a Gaussian noise distribution $p(n(t)) = \mathcal{N}\left(n(t)|\mu_n, \Sigma_n\right)$. Another property of noise is its temporal or spatial correlation. This can be described by correlating the signal with itself, which is called *autocorrelation function*. An extreme case is white noise. White noise is temporally or spatially uncorrelated, meaning the autocorrelation function is a Dirac impulse. The spectrum of white noise is constant, i. e., it contains all frequencies to the same amount as a white light source would contain all visible wavelengths (cf. Fig. 2.8).

2.5 Examples

To conclude this chapter, we want to show the introduced concepts of convolution and Fourier transform on two example systems. A simple system is a smoothing filter, that allows only slow changes of the signal. This is called a low-pass filter. It is an important building block in many applications, for example to remove high-frequency noise from a signal or to remove signal parts with high-frequency before down-sampling to avoid aliasing.

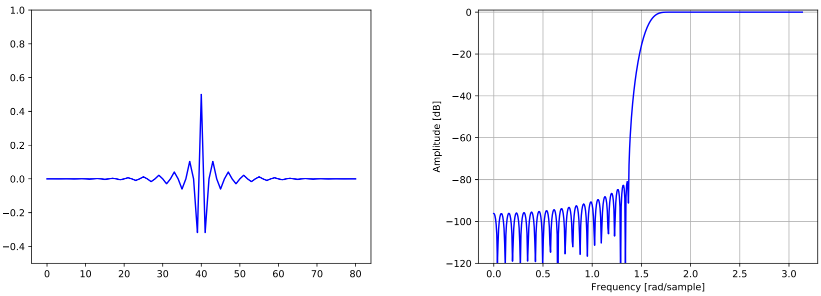
The filter coefficients of a low-pass filter are visualized in Fig. 2.9(a). The low-pass filter has a cutoff frequency of $\frac{\pi}{2}\frac{\text{rad}}{\text{sample}}$ and a length of 81 coefficients. The true properties of the low-pass filter are best perceived in the frequency domain, as displayed in Fig. 2.9(b). Note that the scale of the y-axis is logarithmic. In this context, values of 0 indicate that the signal can pass unaltered. Small values indicate that the signal components are damped. In this example, high frequencies are suppressed by several orders of magnitude. An ideal low-pass filter is a rectangle in the Fourier domain, i. e., all values below the cutoff frequency are passed unaltered and all values above are set



(a) Time domain coefficients

(b) Spectrum

Figure 2.9: Example of a low-pass filter



(a) Time-domain coefficients

(b) Spectrum

Figure 2.10: Example of a high-pass filter

to 0. In our discrete filter, we can only approximate this shape. In the time-domain, the coefficients are samples of a sinc function, which is the inverse Fourier transform of a rectangular function in Fourier domain (cf. Tab. 2.2). The opposite of the low-pass filter is the high-pass filter, shown in Fig. 2.10. Here, frequencies below the cutoff frequency are suppressed, whereas frequencies above are unaltered. Note that the time domain versions of high- and low-pass filters are difficult to differentiate.

Finally, we show how a signal with high and low frequency components is transformed after convolution with a high-pass and a low-pass filter. The signal in Fig. 2.11 is a sine with additive white noise. Thus, noise is distributed equally in the whole frequency domain. A large portion of the noise can be removed by suppressing frequency components where no signal is present. Consequently, the cutoff frequency of the filters is slightly above the frequency of the sine function. As a result, the output of the high-pass filter is similar to the noise and the output of the low-pass filter is similar to the sine. In our

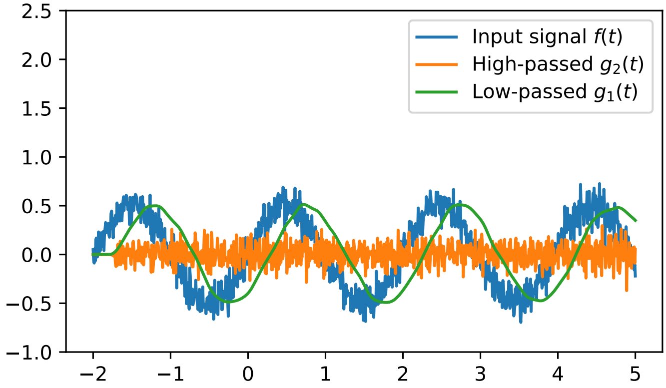


Figure 2.11: Sine signal with additive noise after processing with a low-pass filter and a high-pass filter.

example, we chose a causal filter which introduces a time delay in the filter output. A causal filter can only react to past inputs and needs to collect a certain amount of samples before the filtered result appears at the output.

Further Reading

[1] Ronald N Bracewell. *The Fourier transform and its applications*. McGraw-Hill, New York, 1986.
[2] Olaf Dössel. *Bildgebende Verfahren in der Medizin. Von der Technik zur medizinischen Anwendung*. Vol. 1. 2000.
[3] Bernd Girod, Rudolf Rabenstein, and Alexander Stenger. *Einführung in die Systemtheorie*. Vol. 4. Teubner Stuttgart, 1997.
[4] Andreas Maier. *Speech of children with cleft lip and palate: Automatic assessment*. Logos-Verlag, 2009.
[5] Karl Pearson. "Mathematical contributions to the theory of evolution.—on a form of spurious correlation which may arise when indices are used in the measurement of organs". In: *Proceedings of the royal society of london* 60.359-367 (1897), pp. 489–498.