

Chapter 2

Signals and Systems

The concepts covered in this chapter form the basis for modeling and analyzing communication systems. We start by defining some common signals and studying their properties. In the sequel we address the time and the frequency representation of such signals through the Fourier analysis. The sampling theory is briefly discussed, and linear systems are then put into focus, with emphasis in linear, time-invariant (LTI) systems. The complex representation of signals and systems is also covered and the chapter ends with an analysis of the power spectral density and bandwidth of signals, complementing the study started in Chap. 1.

2.1 Signals

In communications systems, a signal is usually a function of time. In other words, the evolution of a communication signal manifests in the time-domain. Moreover, a communication signal is usually a current or, more frequently, a voltage signal. Signals derived from these two can also arise, though less frequently.

In this section we review the signal classification and present some signals typically encountered in the study of communication systems.

2.1.1 Classification of Signals

There are several ways of classifying a signal. In what follows we present a brief overview of the classifications most frequently used in the study of communication systems.

2.1.1.1 Continuous and Discrete Signals

In a continuous-time signal, the independent variable, *time*, is continuous and the signal is defined continuously in time. Discrete-time signals are defined only at specific time instants. In other words, the independent variable of a discrete-time signal assumes only discrete integer values.

Signals can also be continuous or discrete in their values, leading to the combined classifications: *continuous-times, continuous-valued*; *discrete-time, continuous-valued*; *discrete-time, discrete-valued* or *continuous-time, discrete-valued*. These classifications are illustrated in Fig. 2.1.

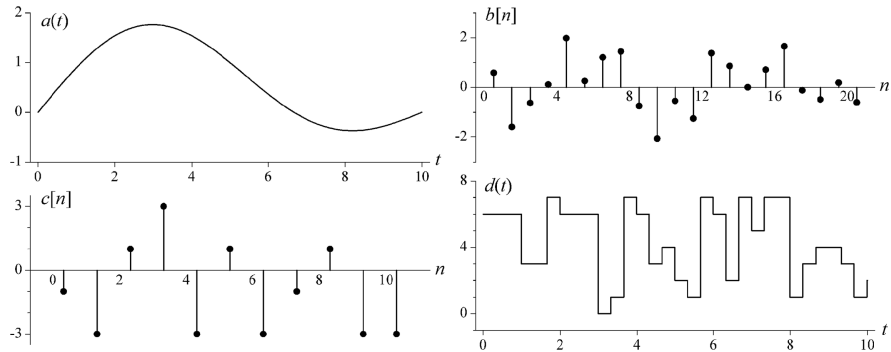


Fig. 2.1 Signal classification: continuous-time, continuous-valued $a(t)$; discrete-time, continuous-valued $b[n]$; discrete-time, discrete-valued $c[n]$ and continuous-time, discrete-valued $d(t)$

Note that, for a discrete-time signal, there is no need to have an explicit association between a value of the discrete time n and a value of the continuous real time t . In practice, this association is made when a discrete-time signal is generated from samples of a continuous-time signal. This association also arises when the result of some signal processing task is used in a real-time application.

Note also that the signal $d(t)$ in Fig. 2.1 is indeed a continuous-time signal, no matter how steep is the transition between the amplitude levels, since in a given instant t there will be a one-to-one correspondence with an amplitude level. This statement comes from the fact that neither a voltage nor a current signal can exhibit a zero transition time, which would correspond to an infinite slew-rate.

2.1.1.2 Periodic and Non-periodic Signals

A continuous-time periodic signal exhibits the same value at integer multiples of T , where T is called the period of the signal. Analogously, a discrete-time periodic signal does not change its value for any discrete-time shift N . Mathematically, for any t and N we have, respectively,

$$\begin{aligned} x(t) &= x(t + T) \quad \text{and} \\ x[n] &= x[n + N]. \end{aligned} \tag{2.1}$$

When (2.1) is not satisfied for a continuous-time and a discrete-time signal, respectively, we say that the signal is non-periodic (or aperiodic).

2.1.1.3 Even and Odd Signals

A continuous-time signal and a discrete-time signal are said to be *even* if they satisfy, respectively,

$$\begin{aligned} x(-t) &= x(t) \quad \text{and} \\ x[-n] &= x[n]. \end{aligned} \quad (2.2)$$

Analogously, a continuous-time signal and a discrete-time signal are said to be *odd* if they satisfy, respectively,

$$\begin{aligned} x(-t) &= -x(t) \quad \text{and} \\ x[-n] &= -x[n]. \end{aligned} \quad (2.3)$$

The classification of even and odd functions can be extended to any domain. We shall also find such classification, for example, when dealing with the frequency content of a communication signal.

2.1.1.4 Deterministic and Random Signals

In a deterministic signal a future value can be precisely predicted. Normally a deterministic signal is described in terms of a deterministic expression, like $x(t) = A \cos(2\pi ft)$, in which for any value of t it is possible to know the value of $x(t)$.

Random signals can not be predicted precisely. A future value or range of values can be predicted only in probabilistic terms, as we have seen in Chap. 1.

When dealing with random communication signals, we normally refer to $x(t)$ as a sample-function of the random process $X(t)$.

2.1.1.5 Baseband and Passband Signals

According to [8, p.60], a baseband signal is “a signal that is not modulated onto a carrier”. As we shall see later on in this chapter, a baseband signal has its frequency content located around $f = 0$, but not necessarily having a non-zero DC component.

A passband signal is a signal whose frequency content is located around a sinusoidal carrier with frequency f_c . Passband signals are commonly associated to the process of modulation in which the information, usually a baseband signal, changes some characteristics of the carrier prior to transmission.

The transmission of information through baseband signals will be discussed in Chap. 4, and in Chap. 6 a number of passband signals generated by means of digital modulations will be studied.

2.1.1.6 Energy and Power Signals

The time-average power of a continuous-time signal is defined by

$$P = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |x(t)|^2 dt. \quad (2.4)$$

For a discrete-time signal we have an analogous definition:

$$P = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N |x[n]|^2. \quad (2.5)$$

In (2.4) and (2.5) the modulus operation takes into account that $x(t)$ or $x[n]$ can be complex signals. Note that if $x(t)$ or $x[n]$ are voltage or current signals, the average power obtained from (2.4) and (2.5) will be measured in watts, normalized to a 1 ohm load resistance.

For periodic signals, (2.4) and (2.5) need to be evaluated only in the interval or number of samples corresponding to the period of the signal.

An *energy signal* has a finite total energy and, as a consequence, has a zero average power. All signals confined in some time interval can be considered energy signals. From (2.4) and (2.5), the energy of a continuous-time and a discrete-time signal confined in $[-T, T]$ or $[-N, N]$ can be determined respectively by

$$\begin{aligned} E &= \int_{-T}^T |x(t)|^2 dt \quad \text{and} \\ E &= \sum_{n=-N}^N |x[n]|^2. \end{aligned} \quad (2.6)$$

As an example, a pulse with amplitude A and duration T is an energy signal with zero power and energy $E = A^2T$.

A *power signal* is a signal that has finite average power and, as a consequence, its energy is infinite. For example, the signal $x(t) = A \cos(2\pi ft)$, $-\infty < t < +\infty$ has average power $P = A^2/2$ and infinite energy.

There are signals that have both P and E infinite, although they are not commonly found in communication systems. As an example we have $x(t) = t$.

2.1.2 Typical Deterministic Signals

In this subsection we present the definitions of some deterministic signals commonly encountered in the study of communication systems.

2.1.2.1 Unit-Step Function

When defined in the time-domain, the continuous-time and discrete-time unit-step functions are given respectively by

$$\begin{aligned} u(t) &= \begin{cases} 1, & t > 0 \\ 0, & t < 0 \end{cases} \quad \text{and} \\ u[n] &= \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases}. \end{aligned} \quad (2.7)$$

The unit-step function is also known as *heaviside step function*. It is defined by some authors also for $t = 0$.

2.1.2.2 Dirac Delta and Kronecker Delta Functions

The first derivative of the continuous-time unit-step function gives rise to the Dirac delta function, as shown by:

$$\delta(t) = \frac{d}{dt}u(t). \quad (2.8)$$

The Dirac delta is a mathematical abstraction consisting of an impulse of infinite amplitude and zero width, and having an area of unity. This abstraction occurs because $u(t)$ is in fact not differentiable, since it is not continuous at $t = 0$. Then we must think of (2.8) as the derivative of a step function with progressive smaller rise time Δ . As Δ decreases, the derivative $\delta_\Delta(t)$ becomes higher in amplitude, smaller in duration and always with unit area. For $\Delta \rightarrow 0$, $\delta_\Delta(t) \rightarrow \delta(t)$.

The *first difference*¹ of the discrete-time unit-step function gives rise to the Kronecker delta function, that is,

$$\delta[n] = u[n] - u[n - 1] = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}. \quad (2.9)$$

Without loss of their particularities, the Dirac and the Kronecker delta functions are usually referred to as *unit impulses*.

The *sifting* property applied to the unit impulses yields, respectively,

$$\begin{aligned} x(t) &= \int_{-\infty}^{\infty} x(\tau)\delta(\tau - t)d\tau \quad \text{and} \\ x[n] &= \sum_{k=-\infty}^{\infty} x[k]\delta[k - n]. \end{aligned} \quad (2.10)$$

The term *sifting* comes from the fact that, in the discrete-time case, the sum in (2.10) goes through the sequence $\{x[k]\}$, retaining or *sifting* the values where $k = n$ as if we had a special *sieve* for these values when $k = n$. The interpretation for the continuous-time case is analogous and follows immediately.

The sifting property can be interpreted alternatively as follows: the integral of a continuous-time function multiplied by a unit impulse results in the value of the function where the impulse exists. An alternative form of the integral in (2.10) illustrates this interpretation:

¹ The *difference* operation on a discrete-time signal is equivalent to the *derivative* operation on a continuous-time signal. Analogously, a *difference equation* in the discrete-time domain corresponds to a *differential equation* in the continuous-time domain.

$$\int_{-\infty}^{\infty} x(t)\delta(t - t_0)dt = x(t_0). \quad (2.11)$$

A similar interpretation follows immediately for the discrete-time case.

The unit impulses play a major role in the characterization of *linear systems*, as we shall see later on in this chapter. Unit impulses are also important in the study of the *sampling theory*, a subject also treated later on in this chapter.

Below we have other properties of the unit impulse function. For the sake of brevity, only the continuous-time Dirac delta function is considered. Similar properties hold for the discrete-time impulse.

$$\begin{aligned} 1. \text{ Scaling : } & \int_{-\infty}^{\infty} \delta(\alpha x)dx = \int_{-\infty}^{\infty} \delta(u)\frac{du}{|\alpha|} = \frac{1}{|\alpha|} \Rightarrow \delta(\alpha x) = \frac{\delta(x)}{|\alpha|} \\ 2. \text{ Generalized scaling : } & \delta[g(x)] = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|} \Rightarrow \delta[\alpha g(x)] = \frac{\delta[g(x)]}{|\alpha|} \\ 3. \text{ Integral/scaling : } & \int_{-\infty}^{\infty} f(x)\delta[g(x)]dx = \sum_i \frac{f(x_i)}{|g'(x_i)|} \\ 4. \text{ Convolution : } & f(t) * \delta(t - T) = f(t - T). \end{aligned} \quad (2.12)$$

Due to the abstraction in the definition of the Dirac delta function, its mathematical treatment becomes difficult. To overcome this difficulty, the delta function is sometimes represented as approximations of other functions. For example, if we take a Gaussian probability density function $f_X(x)$ and make the standard deviation σ progressively small, in the limit as $\sigma \rightarrow 0$, $f_X(x) \rightarrow \delta(t)$. There are several other functions that, in the limit of a given parameter, tend to the unit impulse. See for example [39].

2.1.2.3 Rectangular Function

The rectangular continuous-time and discrete-time functions can be defined respectively by

$$\begin{aligned} \text{rect}\left(\frac{t}{T}\right) &= u\left(t + \frac{T}{2}\right) - u\left(t - \frac{T}{2}\right) = \begin{cases} 1, & |t| \leq T/2 \\ 0, & |t| > T/2 \end{cases} \quad \text{and} \\ \text{rect}\left[\frac{n}{N}\right] &= u\left[n + \frac{N}{2}\right] - u\left[n - \frac{N}{2}\right] = \begin{cases} 1, & |n| \leq N/2 \\ 0, & |n| > N/2 \end{cases}. \end{aligned} \quad (2.13)$$

Although the continuous-time version of the *rect* function can be defined differently from the first equation in (2.13), the definition above is commonly encountered in references on communication systems. On the other hand, it is not so common the association of the term *rect* with the discrete-time version of the rectangular function. In this case it is simply defined by

$$x[n] = \begin{cases} 1, & |n| \leq N \\ 0, & |n| > N \end{cases} . \quad (2.14)$$

2.1.2.4 Sinc Function

The *sinc* function is defined according to

$$\text{sinc}(x) = \text{Sa}(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x}, & x \neq 0 \\ 1, & x = 0 \end{cases} . \quad (2.15)$$

The name *sinc* comes from the function full name: *sine cardinal*. The less usual name *Sa* comes from the word *sample*, as a resemblance of the use of the *sinc* function as an optimum interpolation function to restore an analog signal from its samples. This topic will be treated in the section on sampling theory.

Some authors do not explicitly define the value of $\text{sinc}(x)$ at $x = 0$, since from the L'Hôpital's rule,

$$\lim_{x \rightarrow 0} \frac{\sin(\pi x)}{\pi x} = \lim_{x \rightarrow 0} \frac{\pi \cos(\pi x)}{\pi} = 1. \quad (2.16)$$

Some mathematical software tools like *Mathcad* need the definition of the *sinc* function according to (2.15). For these software tools, $\text{sinc}(x)$ at $x = 0$ is zero. VisSim/Comm assumes that $\text{sinc}(x)$ at $x = 0$ is undefined, since it is interpreted as a division by zero.

Some properties and relations involving the *sinc* function are given below:

$$\begin{aligned} 1. \text{ Infinite product : } \text{sinc}(x) &= \lim_{m \rightarrow \infty} \prod_{n=1}^m \left(1 - \frac{x^2}{n^2}\right) \\ 2. \text{ Gamma function : } \text{sinc}(x) &= \frac{1}{\Gamma(1+x)\Gamma(1-x)} \\ 3. \text{ Dirac delta : } \lim_{a \rightarrow 0} \frac{1}{a} \text{sinc}\left(\frac{x}{a}\right) &= \delta(x). \end{aligned} \quad (2.17)$$

Note in the last line of (2.17) that the *sinc* function is also an approximation or alternative definition for the Dirac delta function.

As mentioned above, the *sinc* function has applications in the sampling theory. It is also one of the most frequently used functions in the study of communication systems, both in the time and in the frequency domains.

2.2 Fourier Analysis of Signals

Jean Baptiste Joseph *Fourier* was a French mathematician and physicist. He was born on March 21, 1768 and died on May 16, 1830. He is known mainly for the application of the so-called Fourier series in the analysis of heat flow.

In fact, the contributions by Joseph Fourier entered into scene in a very controversial discussion around trigonometric sums, that is, the sums of harmonically related sines and cosines and its applications. This subject has gained attention almost a century before Fourier was born, and most of Fourier's discoveries were put into practice only after his lifetime [24].

Today, the broader field of *Fourier analysis* goes beyond the Fourier series and encounters many applications, ranging from physics and signal processing to probability, statistics, random processes and telecommunications, only to mention a few examples.

In this section we review some of the main concepts related to the Fourier analysis of signals, with focus on its applications in digital transmission systems. This review was based on the books by S. Haykin and B. V. Veen [11], H. P. Hsu [13], B. P. Lathi [15–17], and A. V. Oppenheim, A. V. Willsky and S. H. Nawab [24].

Complementary to the analysis of signals, the Fourier analysis of systems will be addressed in Sect. 2.4, where linear systems are put into focus.

2.2.1 Fourier Series

The Fourier series is the representation of a periodic signal by a linear combination of *sines* and *cosines* or *complex exponentials*. When sines and cosines are combined we have the *trigonometric Fourier series*. When complex exponentials are combined we have the *complex exponential Fourier series*.

Both of these complex exponentials or sines and cosines are harmonically related, which means that their frequencies are integer multiples of the *fundamental angular frequency*. For a continuous-time signal, the fundamental angular frequency (in radians per second) is $\omega_0 = 2\pi f_0 = 2\pi/T$, where T is the period of the signal. For a discrete-time signal, the fundamental angular frequency (in radians) is $\Omega_0 = 2\pi f_0 = 2\pi/N$, where N is the period of the signal.

In what follows we consider only the complex exponential form of the Fourier series representation.

2.2.1.1 Fourier Series for Continuous-Time Periodic Signals

For a continuous-time periodic signal, the Fourier series is given by

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk(2\pi/T)t}, \quad (2.18)$$

where T is the period of $x(t)$ and the Fourier series coefficients are determined by

$$a_k = |a_k|e^{j\phi_k} = \frac{1}{T} \int_T x(t)e^{-jk(2\pi/T)t} dt. \quad (2.19)$$

The set of (possibly complex) coefficients $\{a_k\}$ measure the value of each of the frequency components of $x(t)$. The special value of a_0 is the DC component of $x(t)$, that is, the frequency portion of $x(t)$ at $f = 0$.

As we shall see later on in this section, the frequency spectra given by the Fourier series coefficients is plotted as lines or bars in a single graph if $\{a_k\}$ is real or in two graphs if $\{a_k\}$ is complex.

2.2.1.2 Convergence of the Fourier Series for Continuous-Time Periodic Signals

A periodic signal $x(t)$ has a Fourier series representation if it satisfies the Dirichlet conditions:

- It is absolute integrable in one period, that is $\int_T |x(t)|dt < \infty$.
- It has a finite number of maxima and minima and a finite number of finite discontinuities within a period.

Fortunately, a very few signals do not satisfy the Dirichlet conditions, so that the convergence of the Fourier series usually will not be of main concern.

The Fourier series can be truncated so as to consider $2N + 1$ terms:

$$x_N(t) = \sum_{k=-N}^N a_k e^{jk(2\pi/T)t}. \quad (2.20)$$

In this case, if we define an error signal as

$$e_N(t) = x(t) - \sum_{k=-N}^N a_k e^{jk(2\pi/T)t}, \quad (2.21)$$

the convergence of the Fourier series means that as $N \rightarrow \infty$ the energy in the error signal tends to zero, that is,

$$\lim_{N \rightarrow \infty} \int_T |e_N(t)|^2 dt = 0. \quad (2.22)$$

Special attention must be directed to the convergence of the Fourier series for discontinuous signals. The example below aims at clarifying this matter.

Example 2.1 – Consider the square wave $x(t)$ depicted in Fig. 2.2. First let us find the Fourier series coefficients and then analyze the convergence of the Fourier series for a finite number of terms, as determined by (2.20).

From (2.19), the DC level of $x(t)$ is

$$a_0 = \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt = \frac{1}{T} \int_{-T_h/2}^{T_h/2} 1 dt = \frac{T_h}{T}. \quad (2.23)$$

By definition, the *duty-cycle* of a square wave is the ratio between the high level pulse (or mark) duration and the period. Then, the DC value of a square wave signal can be determined by the multiplication of the duty-cycle by the amplitude of the signal, which is a quite obvious result.

The remaining coefficients for $k \neq 0$ can be computed as follows:

$$\begin{aligned} a_k &= \frac{1}{T} \int_{-T/2}^{T/2} x(t) e^{-jk(2\pi/T)t} dt = \frac{1}{T} \int_{-T_h/2}^{T_h/2} e^{-jk(2\pi/T)t} dt \\ &= -\frac{1}{jk2\pi} [e^{-jk(2\pi/T)t}]_{-T_h/2}^{T_h/2} = -\frac{1}{jk2\pi} [e^{-jk\pi(T_h/T)} - e^{jk\pi(T_h/T)}] \\ &= \frac{1}{k\pi} \left[\frac{e^{jk\pi(T_h/T)} - e^{-jk\pi(T_h/T)}}{2j} \right] = \frac{1}{k\pi} \sin \left[k\pi \frac{T_h}{T} \right]. \end{aligned} \quad (2.24)$$

Observe that the set of Fourier series coefficients are real in this case and they will be real in all cases where $x(t)$ is an even function of t . In these cases we need just one graph to depict $\{a_k\}$, as illustrated in Fig. 2.3 for a duty-cycle of 30%.

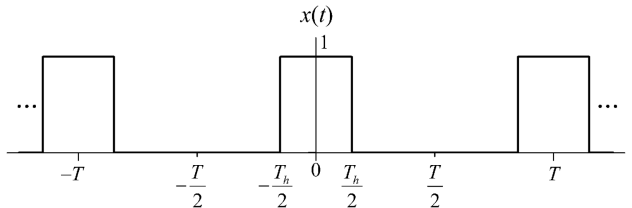


Fig. 2.2 A square wave with duty-cycle $T_h/T\%$

By using a set with $2N + 1$ coefficients in (2.20) we have the so-called *truncated Fourier series approximation*. The synthesis of $x_N(t)$ for some values of N

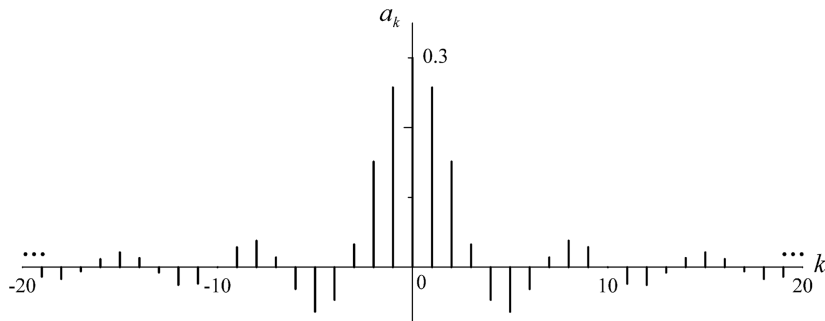


Fig. 2.3 Fourier series coefficients for a square wave with 30% duty-cycle and unit amplitude

are shown in Fig. 2.4. Observe that as N becomes larger, the resultant waveform approaches the original signal $x(t)$. Nevertheless, no matter how large we make N , there will always be the overshoot and undershoot spikes nearby the signal transitions. For a square wave, these peaks have a maximum overshoot and undershoot of about 9% of the signal excursion [17, p. 204; 19, p. 800] and this value does not decrease as N increases. This is known as the *Gibbs phenomenon*, and it will appear whenever the original signal has discontinuous transitions.

The Gibbs phenomenon seems to contradict the convergence properties of the Fourier series, but in fact it does not. As N becomes larger, the overshoots and undershoots remain unchanged in amplitude, but they are shortened in duration. Then, for $N \rightarrow \infty$ the energy of the error signal $e_N(t)$ given in (2.21) tends to zero as it should.

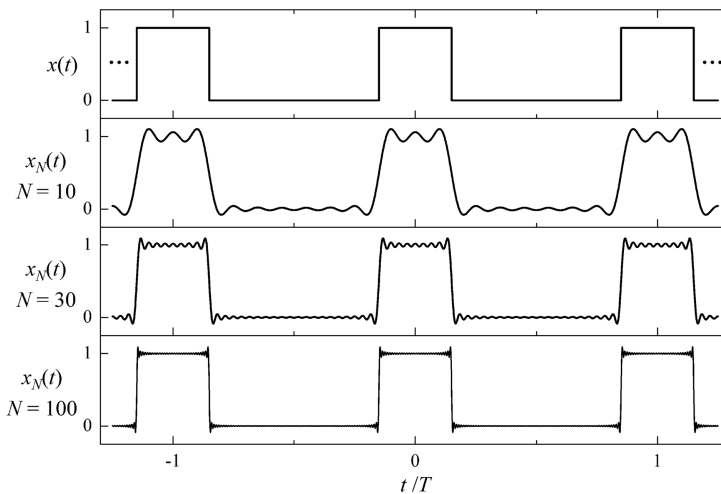


Fig. 2.4 Synthesis of $x_N(t)$ using $2N + 1$ Fourier series coefficients

The reasons for the appearance of the Gibbs phenomenon are briefly discussed in [17, p. 204]. A deeper mathematical treatment on the convergence of the Fourier series and on the Gibbs phenomenon is given in [9].

Simulation 2.1 – Gibbs Phenomenon



File – CD drive: \Simulations\Signals\Gibbs.vsm.

Default simulation settings: Frequency = 10,000 Hz; End = 1 second.
Number of coefficients in the synthesis of $x(t)$ by its truncated Fourier series: $2N + 1 = 33 \Rightarrow N = 16$.

In this experiment we complement the concepts presented above on the Fourier series for continuous-time periodic signals, mainly in what concerns the Gibbs

phenomenon. Three reference signals $x(t)$ are generated, along with their approximate synthesis by the truncated Fourier series $x_N(t)$. These signals are: 1) a single period of a 30% duty-cycle square wave, 2) two periods of a saw-tooth wave and 3) two periods of a triangular wave. All signals have unitary peak-to-peak value.

The energy of the error signal $e_N(t)$, as defined by (2.22), is calculated for each reference and synthesized signals. Additionally, a meter is used to measure the overshoots and undershoots, if any, in order to permit a quantification of the Gibbs phenomenon. Furthermore, the value of N in (2.20) can be adjusted from 1 to 512 in integer powers of 2. The visual similarities between the original and the synthesized signals can be visualized through plots A, B and C.

For each of the signals $x(t)$, vary the value of N and observe the waveforms of $x(t)$ and $x_N(t)$. Note that as the value of N is increased, these waveforms become progressively more similar to one another. Note also that the energy of the error signal reduces progressively as N increases.

Now, while visualizing $x(t)$ and $x_N(t)$ for several values of N , note that the Gibbs phenomenon appears only for the square and the saw-tooth waveforms, a result that is in agreement to what was stated before: the Gibbs phenomenon only happens if the waveform has discontinuities in it.

Pay attention to the overshoot measurements for the square and the saw-tooth waves for several values of N . Note that as soon as the Gibbs phenomenon can be clearly identified, which happens for $N > 8$, the overshoot of $x_N(t)$ referenced to $x(t)$ remains approximately constant in about 9%, confirming the theoretical explanation given before.

The energy of the error signal for the triangular wave, as happens to the other two waveforms, also diminishes as N increases. Since the Gibbs phenomenon does not happen, this energy reduces faster as compared to the error signals produced by the synthesis of the square and the saw-tooth waveforms.

Explore inside the individual blocks. Try to understand how they were implemented. Create and investigate for yourself new situations and configurations of the simulation parameters and try to reach your own conclusions. Specifically, magnify the waveforms near the transitions or peaks for a better view of details. As an exercise, try to explain the operation of the “overshoot meters” used in the simulation.

2.2.1.3 Power of a Continuous-Time Periodic Voltage or Current Signal

The Parseval’s theorem applied to the Fourier series allows for the computation of the average power of a voltage or current waveform $x(t)$ according to

$$P = \frac{1}{T} \int_T |x(t)|^2 dt = \sum_{k=-\infty}^{\infty} |a_k|^2. \quad (2.25)$$

The average power in each harmonic component of $x(t)$ can be determined by

$$P_k = \frac{1}{T} \int_T |a_k|^2 dt = |a_k|^2. \quad (2.26)$$

The above average powers are said to be *normalized*, in the sense that their computation implicitly assumes that the load resistance is 1Ω .

2.2.1.4 Properties of Fourier Series for Continuous-Time Periodic Signals

The most important properties of the Fourier series for continuous-time periodic signals are listed in Table 2.1. For proofs and further details, please refer to [24].

Table 2.1 Properties of the continuous-time Fourier series

Property	Periodic signal	Coefficients
Linearity	$Ax(t) + By(t)$	$Aa_k + Bb_k$
Time shifting	$x(t - \tau)$	$a_k \exp[-jk(2\pi/T)\tau]$
Frequency shift	$x(t) \exp[jM(2\pi/T)t]$	a_{k-M}
Conjugation	$x^*(t)$	a_{-k}^*
Time reversal	$x(-t)$	a_{-k}
Time scaling	$x(\alpha t), \alpha > 0$	a_k
Periodic convolution	$\int_T x(\tau)y(t - \tau)d\tau$	$Ta_k b_k$
Multiplication	$x(t)y(t)$	$\sum_{j=-\infty}^{\infty} a_j b_{k-j}$
Differentiation	$\frac{d}{dt}x(t)$	$jk(2\pi/T)a_k$
Integration	$\int_{-\infty}^t x(t)dt$	$[jk(2\pi/T)]^{-1}a_k$
Symmetry	If $x(t)$ is real, then:	$\begin{cases} a_k = a_{-k}^* \\ \text{Re}\{a_k\} = \text{Re}\{a_{-k}\} \\ \text{Im}\{a_k\} = -\text{Im}\{a_{-k}\} \\ a_k = a_{-k} \\ \arg\{a_k\} = -\arg\{a_{-k}\} \end{cases}$
Real and even signals	If $x(t)$ is real and even, then:	a_k is real and even
Real and odd signals	If $x(t)$ is real and odd, then:	a_k is imaginary and odd
Decomposition of real signals in even and odd parts.	$x_e(t)$ is the even part of $x(t)$: $x_o(t)$ is the odd part of $x(t)$:	$\text{Re}\{a_k\}$ $j\text{Im}\{A_k\}$

2.2.1.5 Fourier Series for Discrete-Time Periodic Signals

For a discrete-time periodic signal, the Fourier series is determined by

$$x[n] = \sum_{k=i}^{i+N-1} a_k e^{jk(2\pi/N)n}, \quad \forall i, \quad (2.27)$$

where N is the period of $x[n]$ and the summation in k from i to $(i + N - 1)$, $\forall i$ means that k starts in any value and varies successively in the range of N samples.

The corresponding Fourier series coefficients are computed from

$$a_k = |a_k|e^{j\phi_k} = \frac{1}{N} \sum_{n=i}^{i+N-1} x[n]e^{-jk(2\pi/N)n}, \forall i. \quad (2.28)$$

Similarly to the continuous-time Fourier series, the set of (possibly complex) coefficients $\{a_k\}$ measure the value of each of the harmonically-related frequency components of $x[n]$. The frequency spectra given by the Fourier series coefficients is plotted as lines or bars in one graph if $\{a_k\}$ is real or in two graphs if $\{a_k\}$ is complex.

An important difference between the Fourier series coefficients for discrete-time aperiodic sequences and for discrete-time periodic sequences is that the former are aperiodic and the later are periodic with period N , that is,

$$a_k = a_{k+N}. \quad (2.29)$$

2.2.1.6 Convergence of the Fourier Series for Discrete-Time Periodic Signals

In contrast to the continuous-time case, convergence aspects of the discrete-time Fourier series are not of concern, since the periodic signal $x[n]$ is precisely synthesized by the linear combination in (2.27), using N harmonically-related complex exponentials and N coefficients determined via (2.28).

2.2.1.7 Power of a Discrete-Time Periodic Voltage or Current Signal

The Parseval's theorem also applies to the calculation of the normalized average power of a voltage or current discrete-time signal $x[n]$ according to

$$P = \frac{1}{N} \sum_{n=i}^{i+N-1} |x[n]|^2 = \sum_{k=i}^{i+N-1} |a_k|^2. \quad (2.30)$$

From (2.30), the average power in the k -th harmonic component of $x[n]$ is $|a_k|^2$.

2.2.1.8 Properties of Fourier Series for Discrete-Time Periodic Signals

The properties of the discrete-time Fourier series are similar to those corresponding to the continuous-time case. Some of these properties are listed in Table 2.2. Again, the reader is referred to [24] for further details and proofs.

2.2.2 Continuous-Time Fourier Transform

In terms of concept, we can interpret the Fourier transform of an aperiodic continuous-time signal $x(t)$ as the limit of its Fourier series for its period tending to infinite. For a better understanding of this concept, revisit Fig. 2.3 in Example 2.1,

Table 2.2 Properties of the discrete-time Fourier series

Property	Periodic signal	Coefficients
Linearity	$Ax[n] + By[n]$	$Aa_k + Bb_k$
Time shifting	$x[n - n_0]$	$a_k \exp[-jk(2\pi/N)n_0]$
Frequency shift	$x[n] \exp[jM(2\pi/N)n]$	a_{k-M}
Conjugation	$x^*[n]$	a_{-k}^*
Time reversal	$x[-n]$	a_{-k}
Time scaling	$x[n/m], n \text{ multiple of } m.$	$(1/m)a_k$
Periodic convolution	$\sum_{r=i}^{i+N-1} x[r]y[n-r], \forall i$	$Na_k b_k$
Multiplication	$x[n]y[n]$	$\sum_{r=i}^{i+N-1} a_r b_{k-r}, \forall i$
First difference	$x[n] - x[n-1]$	$\{1 - \exp[-jk(2\pi/N)]\}a_k$
Running sum	$\sum_{k=-\infty}^n x[k]$	$\{1 - \exp[-jk(2\pi/N)]\}^{-1}a_k$
Symmetry	If $x[n]$ is real, then:	$\begin{cases} a_k = a_{-k}^* \\ \operatorname{Re}\{a_k\} = \operatorname{Re}\{a_{-k}\} \\ \operatorname{Im}\{a_k\} = -\operatorname{Im}\{a_{-k}\} \\ a_k = a_{-k} \\ \arg\{a_k\} = -\arg\{a_{-k}\} \end{cases}$
Real and even signals	If $x[n]$ is real and even, then:	a_k is real and even
Real and odd signals	If $x[n]$ is real and odd, then:	a_k is imaginary and odd
Decomposition of real signals	$x_e[n]$ is the even part of $x[n]$:	$\operatorname{Re}\{a_k\}$
in even and odd parts.	$x_o[n]$ is the odd part of $x[n]$:	$j\operatorname{Im}\{a_k\}$

recalling also that the frequency spacing between any pair of spectral lines is equal to the fundamental ordinary frequency $f_0 = 1/T$ hertz.

Let us rewrite expression (2.24) for these Fourier series coefficients:

$$a_k = \frac{1}{k\pi} \sin\left(k\pi \frac{T_h}{T}\right) = \frac{T_h}{T} \frac{\sin[k\pi(T_h/T)]}{k\pi(T_h/T)} = \frac{T_h}{T} \operatorname{sinc}\left(k \frac{T_h}{T}\right). \quad (2.31)$$

Note that the first null coefficient, if any, will happen for

$$k\pi \frac{T_h}{T} = \pi \Rightarrow k = \frac{T}{T_h}. \quad (2.32)$$

This value of k corresponds to the frequency

$$f_k = k \frac{1}{T} = \frac{T}{T_h} \frac{1}{T} = \frac{1}{T_h}, \quad (2.33)$$

which is independent of T . A continuous function of f that has the same shape of the *envelope* of (2.31) is the *sinc* function

$$X(f) = T_h \operatorname{sinc}(fT_h). \quad (2.34)$$

This function is plotted in Fig. 2.5, along with the Fourier series coefficients of Example 2.1, multiplied by T . This multiplication is made to guarantee that the continuous envelope is kept attached to the values of the coefficients for different values of T . Note in this figure that the horizontal axis was modified to permit the matching between the index of a coefficient and its ordinary frequency.

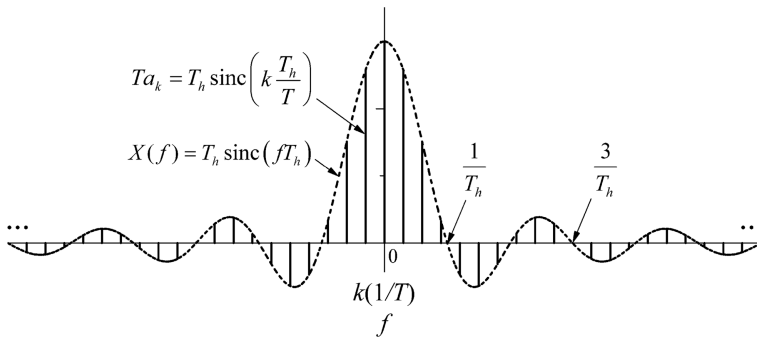


Fig. 2.5 Fourier series coefficients and its envelope for a square wave

If we let T increase in Fig. 2.2, keeping T_h unchanged, the spectral lines in Fig. 2.5 becomes closer. In the limit of $T \rightarrow \infty$, the square wave $x(t)$ will tend to a single rectangular pulse of duration T_h and the modified Fourier series coefficients $T a_k$ will tend to the function $X(f)$. As we shall see later, (2.34) is indeed the Fourier transform of a unit amplitude rectangular pulse of duration T_h and centered about the origin (see Table 2.4 a littler further ahead).

This is the reasoning behind the Fourier transform: we create an aperiodic signal by letting the period of a periodic signal tend to infinity. Then we apply a modified version of the Fourier series, as we shall see in the sequel.

The continuous-time Fourier transform plays a major role in the study of communication systems and, for this reason, a special attention is devoted to it here.

2.2.2.1 Definition of the Continuous-Time Fourier Transform

Let us apply the reasoning constructed above to the expressions of the Fourier series for periodic signals. We shall use $x_p(t)$ to identify the periodic signal and $x(t)$ to identify the aperiodic signal constructed from $x_p(t)$ by letting $T \rightarrow \infty$. From (2.19), the value of $T a_k$ is given by

$$T a_k = \int_T x_p(t) e^{-jk(2\pi/T)t} dt. \quad (2.35)$$

When $T \rightarrow \infty$, $T a_k$ will tend to the function $X(f)$, $x_p(t)$ becomes $x(t)$ and $k(2\pi/T)$ becomes $2\pi f$. With these results in (2.35) we obtain

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt. \quad (2.36)$$

From (2.18), with $a_k = X(f)/T$ and $f_0 = 1/T$ we have

$$x_p(t) = \sum_{k=-\infty}^{\infty} X(f)e^{jk(2\pi/T)t} f_0. \quad (2.37)$$

When $T \rightarrow \infty$, $x_p(t)$ becomes $x(t)$, f_0 will tend to zero, $k(2\pi/T)$ becomes $2\pi f$ and the summation in (2.37) becomes an integral. As a result we obtain

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} df. \quad (2.38)$$

Equations (2.36) and (2.38) form the *Fourier transform pair* of a continuous-time aperiodic signal. The frequency content of $x(t)$ is obtained via (2.36), which is called simply the *Fourier transform* of $x(t)$. The time function $x(t)$ is recovered from its Fourier transform by applying the *inverse Fourier transform* (2.38). We shall denote the Fourier transform of $x(t)$ as $\mathfrak{F}\{x(t)\}$.

Note that if $x(t)$ is a voltage signal, then the Fourier transform will give the voltage of the frequency components of $x(t)$ through a continuum function of the ordinary frequency f .

2.2.2.2 Convergence of the Continuous-Time Fourier Transform

The sufficient conditions for the convergence of the Fourier transform of a continuous-time aperiodic signal $x(t)$ are similar to those presented for the Fourier series. They are also called Dirichlet conditions:

- $x(t)$ must be absolute integrable, that is, $\int_{-\infty}^{\infty} |x(t)|dt < \infty$.
- It must have a finite number of maxima and minima and a finite number of finite discontinuities within a finite interval.

2.2.2.3 Continuous-Time Fourier Transform for Periodic Signals

The Fourier transform of a continuous-time periodic signal $x(t)$ can be obtained from its Fourier series coefficients according to

$$X(f) = \sum_{k=-\infty}^{\infty} a_k \delta\left(f - \frac{k}{T}\right). \quad (2.39)$$

This result means that the Fourier transform of a periodic signal is a train of impulses located in multiples of the fundamental frequency $f_0 = 1/T$ Hz, and whose areas are equal to the values of the Fourier series coefficients.

The values of the coefficients are, as before,

$$a_k = |a_k|e^{j\phi_k} = \frac{1}{T} \int_T x(t)e^{-jk(2\pi/T)t} dt. \quad (2.40)$$

The signal $x(t)$ can be recovered from its Fourier transform by applying

$$\begin{aligned} x(t) &= \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} df \\ &= \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_k \delta\left(f - \frac{k}{T}\right) e^{j2\pi ft} df \\ &= \sum_{k=-\infty}^{\infty} a_k \int_{-\infty}^{\infty} \delta\left(f - \frac{k}{T}\right) e^{j2\pi ft} df \\ &= \sum_{k=-\infty}^{\infty} a_k e^{jk(2\pi/T)t}, \end{aligned} \quad (2.41)$$

where the last line was determined from the preceding line by applying the sifting property of the Dirac delta function.

2.2.2.4 Energy of a Continuous-Time Aperiodic Voltage or Current Signal

The Parseval's theorem applied to the Fourier transform allows for the computation of the normalized average energy of a voltage or current waveform $x(t)$ according to

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df. \quad (2.42)$$

The function $|X(f)|^2$ is called the *energy spectral density* of $x(t)$. It shows how the energy of $x(t)$ is distributed in a continuum of the ordinary frequency f .

2.2.2.5 Properties and Pairs of the Continuous-Time Fourier Transform

Tables 2.3 and 2.4 present some Fourier transform properties and pairs, respectively. For additional details and proofs, please refer to [11, 13, 17, 24].

Table 2.3 Properties of the Fourier transform

Property	Aperiodic signal	Fourier transform
Linearity	$Ax(t) + By(t)$	$AX(f) + BY(f)$
Time shifting	$x(t - \tau)$	$X(f) \exp[-j2\pi f\tau]$
Frequency shift	$x(t) \exp[j2\pi f_0 t]$	$X(f - f_0)$
Conjugation	$x^*(t)$	$X^*(-f)$
Duality	$X(t)$	$x(-f)$
Time reversal	$x(-t)$	$X(-f)$
Time scaling	$x(\alpha t)$	$\frac{1}{ \alpha } X\left(\frac{f}{\alpha}\right)$
Convolution	$x(t) * y(t)$	$X(f)Y(f)$
Multiplication	$x(t)y(t)$	$X(f) * Y(f)$

Table 2.3 (continued)

Property	Aperiodic signal	Fourier transform
Differentiation in time	$\frac{d^n}{dt^n} x(t)$	$(j2\pi f)^n X(f)$
Differentiation in frequency	$t^n x(t)$	$\left(\frac{j}{2\pi}\right)^n \frac{d^n X(f)}{df^n}$
Integration	$\int_{-\infty}^t x(u) du$	$\frac{1}{j2\pi f} X(f) + \frac{X(0)}{2} \delta(f)$
Area under $x(t)$	$\int_{-\infty}^{\infty} x(t) dt$	$X(0)$
Area under $X(f)$	$\int_{-\infty}^{\infty} X(f) df$	$g(0)$
Symmetry	If $x(t)$ is real, then:	$\begin{cases} X(f) = X^*(-f) \\ \operatorname{Re}\{X(f)\} = \operatorname{Re}\{X(-f)\} \\ \operatorname{Im}\{X(f)\} = -\operatorname{Im}\{X(-f)\} \\ X(f) = X(-f) \\ \arg\{X(f)\} = -\arg\{X(-f)\} \end{cases}$
Real and even signals	If $x(t)$ is real and even, then:	$X(f)$ is real and even
Real and odd signals	If $x(t)$ is real and odd, then:	$X(f)$ is imaginary and odd
Decomposition of real signals in even and odd parts.	$x_e(t)$ is the even part of $x(t)$: $x_o(t)$ is the odd part of $x(t)$:	$\operatorname{Re}\{X(f)\}$ $j\operatorname{Im}\{X(f)\}$

Table 2.4 Continuous-time Fourier transform pairs

Signal	Fourier transform
$\operatorname{Arect}\left(\frac{t}{T}\right) = \begin{cases} A, & t < T/2 \\ 0, & t \geq T/2 \end{cases}$	$AT \frac{\sin(\pi f T)}{(\pi f T)} = AT \operatorname{sinc}(fT)$
$\begin{cases} 1 - \frac{ t }{T}, & t < T \\ 0, & t \geq T \end{cases}$	$T \frac{\sin^2(\pi f T)}{(\pi f T)^2} = T \operatorname{sinc}^2(fT)$
$\operatorname{sinc}(2BT)$	$\frac{1}{2B} \operatorname{rect}\left(\frac{f}{2B}\right)$
$\exp(-\alpha t^2), \operatorname{Re}(\alpha) > 0$	$\sqrt{\frac{\pi}{\alpha}} \exp\left[-\frac{(\pi f)^2}{\alpha}\right]$
$\exp(-\alpha t)u(t), \alpha > 0$	$\frac{1}{\alpha + j2\pi f}$
$\exp(j\alpha t)$	$\delta\left(f - \frac{\alpha}{2\pi}\right)$
$u(t)$ (unit step function)	$\frac{1}{2} \left(\frac{1}{j\pi f} + \delta(f) \right)$
$\exp(-\alpha t), \alpha > 0$	$\frac{\alpha^2 + 4\pi^2 f^2}{-j \operatorname{sgn}(f)},$ where $\operatorname{sgn}(f) = u(f) - u(-f)$
1	$\delta(f)$
$\delta(t)$	1
$\cos(\alpha t)$	$\frac{1}{2} \left[\delta\left(f - \frac{\alpha}{2\pi}\right) + \delta\left(f + \frac{\alpha}{2\pi}\right) \right]$
$\sin(\alpha t)$	$\frac{1}{2j} \left[\delta\left(f - \frac{\alpha}{2\pi}\right) - \delta\left(f + \frac{\alpha}{2\pi}\right) \right]$
$\cos(\alpha t^2)$	$\sqrt{\frac{\pi}{\alpha}} \cos\left(\frac{\pi^2 f^2}{\alpha} - \frac{\pi}{4}\right)$
$\sin(\alpha t^2)$	$-\sqrt{\frac{\pi}{\alpha}} \sin\left(\frac{\pi^2 f^2}{\alpha} - \frac{\pi}{4}\right)$
$\sum_{n=-\infty}^{\infty} \delta(t - nT)$	$\frac{1}{T} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T}\right)$

2.2.3 Discrete-Time Fourier Transform

The discrete-time Fourier transform also plays an important role in the study of communication systems. For this reason, special attention is devoted to it here.

2.2.3.1 Definition of the Discrete-Time Fourier Transform

Let $x[n]$ be an aperiodic discrete-time sequence and $X(\Omega)$ its Fourier transform, where Ω is the angular frequency, measured in radians. These functions are related thorough the *discrete-time Fourier transform pair*:

$$\begin{aligned} X(\Omega) &= \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n} \quad \text{and} \\ x[n] &= \frac{1}{2\pi} \int_{2\pi} X(\Omega)e^{j\Omega n} d\Omega. \end{aligned} \quad (2.43)$$

The upper expression in (2.43) is the analysis equation and the lower is the synthesis equation. Note that these equations are the counterparts of (2.36) and (2.38), respectively, with some remarkable differences:

- First, the angular frequency Ω is measured in radians, in contrast to the ordinary angular frequency $\omega = 2\pi f$, measured in radians per second. This is because no time measurement is attached to the discrete-time signal $x[n]$.
- Secondly, note that the integration in (2.43) is limited to the angular frequency interval of 2π . This happens because $X(\Omega)$ is periodic with period 2π , as we shall see in what follows.

Later on in this chapter we shall map the angular frequency Ω into the ordinary frequency f , by attaching a time-scale to the sequence $x[n]$.

2.2.3.2 Periodicity of the Discrete-Time Fourier Transform

The discrete-time Fourier transform is periodic with period 2π , as shown by:

$$\begin{aligned} X(\Omega + 2\pi) &= \sum_{n=-\infty}^{\infty} x[n]e^{-j(\Omega+2\pi)n} = e^{-j2\pi n} \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n} \\ &= \left[\underbrace{\cos(2\pi n)}_1 - j \underbrace{\sin(2\pi n)}_0 \right] \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n} \\ &= \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n} = X(\Omega). \end{aligned} \quad (2.44)$$

2.2.3.3 Convergence of the Discrete-Time Fourier Transform

The convergence of $X(\Omega)$ in (2.43) demands that $x[n]$ is absolutely summable, i.e.,

$$\sum_{n=-\infty}^{\infty} |x[n]| < \infty. \quad (2.45)$$

The expression of synthesis in (2.43) is not of concern in terms of convergence, since the integration interval is finite.

2.2.3.4 Discrete-Time Fourier Transform for Periodic Signals

Let $x[n]$ be a discrete-time periodic sequence with period N . Its Fourier series is given by

$$x[n] = \sum_{k=i}^{i+N-1} a_k e^{jk(2\pi/N)n}, \quad \forall i. \quad (2.46)$$

The Fourier series coefficients are known to be

$$a_k = \frac{1}{N} \sum_{n=i}^{i+N-1} x[n] e^{-jk(2\pi/N)n}, \quad \forall i. \quad (2.47)$$

The discrete-time Fourier transform of the periodic sequence $x[n]$ can be determined from the knowledge of the Fourier series coefficients according to

$$X(\Omega) = \sum_{k=-\infty}^{\infty} 2\pi a_k \delta\left(\Omega - \frac{2\pi k}{N}\right). \quad (2.48)$$

2.2.3.5 Energy of a Discrete-Time Aperiodic Voltage or Current Signal

The Parseval's theorem applied to the discrete-time Fourier transform allows for the computation of the normalized average energy of an aperiodic voltage or current discrete-time signal $x[n]$ according to

$$E = \sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi} \int_{2\pi} |X(\Omega)|^2 d\Omega. \quad (2.49)$$

The function $|X(\Omega)|^2$ is the *energy spectral density* of $x[n]$. It shows how the energy of $x[n]$ is distributed in a continuum of values of the angular frequency Ω .

2.2.3.6 Properties and Pairs of the Discrete-Time Fourier Transform

In what follows we list some of the properties and pairs of the discrete-time Fourier transform. Again, the reader is referred to [24] for further details and proofs.

Table 2.5 Properties of the discrete-time Fourier transform

Property	Aperiodic signal	Fourier transform
Linearity	$\alpha x[n] + \beta y[n]$	$\alpha X(\Omega) + \beta Y(\Omega)$
Time shifting	$x[n - n_0]$	$X(\Omega) \exp[-j\Omega n_0]$
Frequency shift	$x[n] \exp[j\Omega_0 n]$	$X(\Omega - \Omega_0)$
Conjugation	$x^*[n]$	$X^*(-\Omega)$
Time reversal	$x[-n]$	$X(-\Omega)$
Time scaling	$x[n/m]$, n multiple of m .	$X(m\Omega)$
Convolution	$x[n] * y[n]$	$X(\Omega)Y(\Omega)$
Multiplication	$x[n]y[n]$	$\frac{1}{2\pi} \int_{2\pi} X(\theta)Y(\Omega - \theta)d\theta$
First time difference	$x[n] - x[n - 1]$	$[1 - \exp(-j\Omega)]X(\Omega)$
Running sum	$\sum_{k=-\infty}^n x[k]$	$[1 - \exp(-j\Omega)]^{-1} X(\Omega) + \pi X(0)\delta(\Omega)$, $ \Omega \leq \pi$
Differentiation in frequency	$nx[n]$	$j \frac{dX(\Omega)}{d\Omega}$
Symmetry	If $x[n]$ is real, then:	$\begin{cases} X(\Omega) = X^*(-\Omega) \\ \operatorname{Re}\{X(\Omega)\} = \operatorname{Re}\{X(-\Omega)\} \\ \operatorname{Im}\{X(\Omega)\} = -\operatorname{Im}\{X(-\Omega)\} \\ X(\Omega) = X(-\Omega) \\ \arg\{X(\Omega)\} = -\arg\{X(-\Omega)\} \end{cases}$
Real and even signals	If $x[n]$ is real and even, then:	$X(\Omega)$ is real and even
Real and odd signals	If $x[n]$ is real and odd, then:	$X(\Omega)$ is imaginary and odd
Decomposition of real signals in even and odd parts.	$x_e[n]$ is the even part of $x[n]$: $x_o[n]$ is the odd part of $x[n]$:	$\operatorname{Re}\{X(\Omega)\}$ $j\operatorname{Im}\{X(\Omega)\}$
Periodicity	$x[n]$	$X[\Omega] = X[\Omega + 2\pi]$

2.2.4 Discrete Fourier Transform

The discrete Fourier transform (DFT) of a finite discrete-time sequence $x[n]$ produces a finite discrete-frequency sequence $X[k]$ according to

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-jkn(2\pi/N)} = \sum_{n=0}^{N-1} x[n] W_N^{kn}, \quad k = 0, 1, \dots, N-1, \quad (2.50)$$

where W_N is a N -th root of unity² and is defined by

$$W_N = e^{-j(2\pi/N)}. \quad (2.51)$$

² In mathematics, the N -th roots of unity are all the complex numbers that yield 1 when raised to the power N .

Table 2.6 Discrete-time Fourier transform pairs

Signal	Fourier transform
$x[n] = \begin{cases} 1, & n \leq N_1 \\ 0, & n > N_1 \end{cases}$	$\frac{\sin[\Omega(N_1 + \frac{1}{2})]}{\sin(\Omega/2)}$
$\frac{\sin(Wn)}{\pi n}, 0 < W < \pi$	$\begin{cases} 1, & 0 \leq \Omega \leq W \\ 0, & W < \Omega \leq \pi \end{cases}$
$\delta[n]$	1
1	$2\pi\delta(\Omega), \Omega \leq \pi$
$\delta[n - n_0]$	$\exp(-j\Omega n_0)$
$\exp(j\Omega_0 n)$	$2\pi\delta(\Omega - \Omega_0), \Omega , \Omega_0 \leq \pi$
$\cos(\Omega_0 n)$	$\pi[\delta(\Omega - \Omega_0) + \delta(\Omega + \Omega_0)], \Omega , \Omega_0 \leq \pi$
$\sin(\Omega_0 n)$	$-j\pi[\delta(\Omega - \Omega_0) - \delta(\Omega + \Omega_0)], \Omega , \Omega_0 \leq \pi$
$u[n]$ (unit step function)	$\pi\delta(\Omega) + \frac{1}{1 - \exp(-j\Omega)}, \Omega \leq \pi$
$-u[-n - 1]$	$-\pi\delta(\Omega) + \frac{1}{1 - \exp(-j\Omega)}, \Omega \leq \pi$
$a^n u[n], a < 1$	$\frac{1}{1 - a \exp(-j\Omega)}$
$-a^n u[-n - 1], a > 1$	$\frac{1}{1 - a \exp(-j\Omega)}$
$(n + 1)a^n u[n], a < 1$	$\frac{[1 - a \exp(-j\Omega)]^2}{1 - a^2}$
$a^{ n }, a < 1$	$\frac{1 - 2a \cos \Omega + a^2}{1 - 2a \cos \Omega + a^2}$
$\sum_{k=-\infty}^{\infty} \delta[n - kN_0]$	$\frac{2\pi}{N_0} \sum_{k=-\infty}^{\infty} \delta\left(\Omega - k\frac{2\pi}{N_0}\right)$

The inverse discrete Fourier transform (IDFT) is given by

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{jkn(2\pi/N)} = \frac{1}{N} \sum_{k=0}^{N-1} X[k] W_N^{-kn}, \quad n = 0, 1, \dots, N-1. \quad (2.52)$$

Note that the discrete Fourier transform produces a discrete-frequency result, as opposed to the continuous-frequency result produced by the discrete-time Fourier transform. The DFT of a sequence $x[n]$ with length N is usually referred to as an N -point DFT. Similarly, the IDFT of a sequence $X[k]$ with length N is usually referred to as an N -point IDFT.

The DFT and IDFT are very attractive Fourier representations of a signal since they are both finite and, being discrete, they can be operated naturally by computers. In fact, the DFT and the IDFT have very fast algorithms for their computation: the FFT (fast Fourier transform) and IFFT (inverse fast Fourier transform). Furthermore, the DFT and IDFT have a direct relationship with the discrete-time Fourier series and the Fourier transform. We shall briefly discuss these concepts in the sequel.

2.2.4.1 The DFT and the Discrete-Time Fourier Series

Comparing (2.50) and (2.52) with (2.27) and (2.28) we notice that the values of $X[k]$ for a finite sequence $x[n]$ can be determined from the Fourier series coefficients of the periodic extension of $x[n]$, with period N , according to

$$X[k] = Na_k. \quad (2.53)$$

2.2.4.2 The DFT and the Fourier Transform

Comparing the Fourier transform in (2.43) with (2.50) we notice that the DFT $X[k]$ corresponds to samples of $X(\Omega)$ spaced by $2\pi/N$ radians apart, that is,

$$X[k] = X(\Omega)|_{\Omega=k\frac{2\pi}{N}} = X\left(k\frac{2\pi}{N}\right). \quad (2.54)$$

2.2.4.3 Mapping the Discrete-Frequency Index k into the Ordinary Frequency f

As we shall see in the study of linear systems, the angular frequency Ω (in radians) is mapped into the ordinary frequency f (in hertz) according to

$$f = \frac{\Omega}{2\pi T_s}, \quad (2.55)$$

where T_s is the time interval between the samples in the sequence $x[n]$.

Then, a value of the DFT in the discrete-frequency k can be mapped into a frequency f_k (in hertz) by combining (2.55) and (2.54), yielding

$$f_k = \frac{\Omega_k}{2\pi T_s} = \frac{(k2\pi/N)}{2\pi T_s} = \frac{k}{NT_s}. \quad (2.56)$$

2.2.4.4 Parseval's Relation for the DFT

The Parseval's relation applied to the DFT can be written as

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2. \quad (2.57)$$

If $x[n]$ is a discrete-time voltage or current signal, the Parseval's relation allows for the computation of the normalized average energy of $x[n]$ in the time or in the frequency domain. By dividing both sides of (2.57) per N , an estimate of the average power in the discrete-time signal $x[n]$ is obtained. Nevertheless, for aperiodic sequences with large lengths, the precision of the result will be governed by the

value of N : the greater the value, the better the precision of the estimated average power will be.

2.2.4.5 Direct and Inverse Numerical Fast Fourier Transforms (FFT and IFFT)

The numerical computations of the DFT and the IDFT of a finite sequence are usually referred to as fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) algorithms.

The FFT algorithm was motivated by the fact that a direct evaluation of the sums in (2.50) would lead to arithmetic operations with complexity $O(N^2)$.³ Some FFT algorithms compute (2.50) with complexity $O(N \log_2 N)$.

One of the most common FFT algorithms is the Cooley-Tukey [5]. It recursively divides an N -point DFT into DFTs of smaller sizes, an approach known as divide-and-conquer. Although the development of this algorithm is credited to J. W. Cooley and J. W. Tukey, it was later discovered that those authors have independently reinvented the algorithm already known to Carl Friedrich Gauss around 1805 [39, 40].

The most used form of the Cooley-Tukey algorithm divides the N -point DFT into pieces of size $N/2$ at each step. For this reason this form of the algorithm is limited to power-of-two sizes, but other sizes can be used in general.

A detailed analysis about the particular structures of FFT and IFFT algorithms is beyond the scope of this book. The interested reader can obtain a formal treatment on the general aspects of the computation of the DFT in Chap. 9 of [23]. Several FFT algorithms are discussed in [4].

It is worth mentioning that VisSim/Comm makes an extensive use of FFT algorithms to internally compute the frequency content of the signals in a given diagram. Moreover, VisSim/Comm computes the mapping between the discrete-frequency index k to the ordinary frequency f by taking into account the size of the FFT, N , and the simulation frequency, $f_s = 1/T_s$, according to (2.56). The result is then presented in the ordinary frequency f , in the range $[0, f_s/2]$ Hz or $[-f_s/2$ to $f_s/2]$ Hz, depending on the specifics of the frequency plot.

Simulation 2.2 – The FFT via VisSim/Comm



File – CD drive: \Simulations\Signals\FFT.vsm.

Default simulation settings: Frequency = 10,000 Hz; End = 0.4 second. Sinusoidal $x[n]$ period: $N_0 = 20$. Input signal selector: Random $x[n]$.

³ The “big O” notation is commonly used to identify the complexity of algorithms in terms of computational time or memory resources used. A complexity $O(N^2)$ means that the computational resources grows with the square of N or, alternatively, means that N squared dominates the complexity aspects.

This experiment aims at illustrating a realistic computation of the FFT for a periodic and an aperiodic sequence $x[n]$ with length $N_T = 4,000$, as performed by the simulation software VisSim/Comm.

The aperiodic sequence was generated from samples of a band-limited voice sample-function and the periodic signal is a sinusoidal sequence with configurable period. One of these signals is selected by the user for analysis.

The selected signal first goes to average power estimation blocks. One of these blocks belongs to the VisSim/Comm block set and computes the running power estimation

$$P_k = \frac{1}{k} \sum_{n=0}^{k-1} |x[n]|^2. \quad (2.58)$$

At the end of the simulation, the value of P_k produced by (2.58) for $k = N_T$ will correspond to the result obtained by the left-side of (2.57), divided by N_T .

The other average power estimation is an explicit realization of the left-side of (2.57), the result being divided by N_T .

In the lower part of the experiment the selected signal goes through an FFT computation block. This block is performing an FFT with $N = 2,048$ points (2k-point FFT) of $x[n]$. Since the input data has $N_T = 4,000$ points in the simulation interval, the remaining $(N_T - N)$ points are not used for the FFT computation. Note that if we had $N_T = 4,096$ points, a 4k-point FFT would be possible, yielding a more precise estimation of the frequency content of $x[n]$.

The transform calculated by the FFT block corresponds to (2.50). Then, to obtain the correct frequency content of $x[n]$ we must divide the values of $X[k]$ by N , according to (2.53). The result is plotted in linear and logarithmic scales.

The values of the FFT result $X[k]$ are exported to the file “c:\FFTdata.dat”, so that you are able to use it in another application or experiment as desired.

The average power of $x[n]$ is also estimated in the frequency domain by applying the right-side of (2.57), divided by N . Running the simulation you can verify that all three average power estimations are approximately equal to one another.

Run the simulation while observing Plot A. Note that the FFT block gives a bilateral FFT result in the ordinary frequency f . In other words, VisSim/Comm computes the mapping between the discrete-frequency index k to the ordinary frequency f by taking into account the size of the FFT, N , and the simulation frequency, $f_s = 1/T_s$, according to (2.56). The result is then presented in the ordinary frequency f in the range $[-f_s/2$ to $f_s/2]$ Hz. Note also that the magnitude frequency spectrum is an even function of f , since $x[n]$ is a real sequence.

Now, in the “selector” block, select the sinusoidal $x[n]$ and, while observing Plot A, run the simulation. Since the sinusoid has a default period $N_0 = 20$, its angular frequency is $\Omega_0 = 2\pi/N_0$ radians and its ordinary frequency, as determined via (2.55) is $f_0 = \Omega_0/(2\pi T_s) = 500$ Hz. The amplitude of each discrete frequency component will be 0.5, since the amplitude of the sinusoidal sequence is unitary. However, observe that this amplitude is below 0.4, corresponding to a large error. This error was in fact expected, since a 2k-point FFT is being performed in a

periodic signal with period $N_0 = 20$. As a result, we do not have an integer number of periods in the summation (2.50), which produces the error.

The preceding paragraph is intended to serve as a warning: care must be taken when analyzing the frequency content of a periodic signal via FFT. In practical situations, increasing the FFT size, if possible, will reduce the error. Nevertheless, the error will be completely eliminated if N_0 divides N . To confirm this, change the period of the sinusoidal sequence to $N_0 = 16$. Now the summation in (2.50) will consider an integer number of 2, $048/16 = 128$ periods of $x[n]$. Run the simulation and observe Plot A. Note that the magnitude of the discrete frequency components, now located in $f_0 = \Omega_0/(2\pi T_s) = 625$ Hz, is exactly 0.5, as expected.

In fact, only one period of a periodic signal suffices for the DFT computation. However, when the signal is composed by several spectral lines, an increased N will increase the frequency resolution of the frequency plot.

The logarithmic values in Plot A were determined by multiplying the logarithm to base 10 of the original amplitude value by 20.

Explore inside the individual blocks. Try to understand how they were implemented. Create and investigate for yourself new situations and configurations of the simulation parameters and try to reach your own conclusions.

2.2.5 Laplace and Z-Transforms

The Laplace transform [17, p. 361] extends the capabilities of the Fourier transform by considering not only the imaginary axis (as Fourier transform does), but also the real axis. It has a wide range of applications, from circuit analysis to the solution of differential equations and the analysis of feedback systems. In communications, its major application is in the analysis of linear systems [13, p. 110; 24, p. 573; 11, p. 401].

The Z-transform [17, p. 669], likewise the discrete-time Fourier transform, converts a discrete time-domain signal into a complex frequency-domain representation. Nevertheless, Z-transform is capable of extending the capabilities of the discrete-time Fourier transform. Its main application is in the digital (discrete-time) signal processing [23, p. 149] area.

In what follows we present some brief comments restricted to the correspondence between the Laplace and the Fourier transform and between the Z-transform and the discrete-time Fourier transform. The interested reader is encouraged to consult the references cited throughout this subsection for further details.

The *bilateral Laplace transform* of a continuous-time signal $x(t)$ is given by

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st} dt. \quad (2.59)$$

Recalling the continuous-time Fourier transform of $x(t)$, which is given by

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt, \quad (2.60)$$

we notice that it is a special case of the Laplace transform in which $s = j2\pi f$. Then we can write

$$X(s)|_{s=j2\pi f} = \Im\{x(t)\} = X(f). \quad (2.61)$$

Since the algebraic expression for the Laplace transform of a given function can be equal to the expression related to a different function, the complete specification of the Laplace transform demands the determination of the range of values of s for which the transform converges. This range of values is called the *region of convergence* (ROC) of the Laplace transform. As an example, consider the Laplace transform of $x(t)$ and $y(t)$, as shown below [13, p. 111]:

$$\begin{aligned} 1. x(t) = e^{at}u(t), a \text{ real} &\Rightarrow X(s) = \int_{-\infty}^{\infty} x(t)e^{-st} dt = \frac{1}{s+a}, \operatorname{Re}(s) > -a \\ 2. y(t) = -e^{-at}u(-t), a \text{ real} &\Rightarrow Y(s) = \int_{-\infty}^{\infty} y(t)e^{-st} dt = \frac{1}{s+a}, \operatorname{Re}(s) < -a. \end{aligned} \quad (2.62)$$

Observe that the algebraic expressions of the Laplace transform of $x(t)$ and $y(t)$ are equal, but they differ in their ROC.

The *bilateral Z-transform* of the discrete-time sequence $x[n]$ is defined by

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}. \quad (2.63)$$

Now, recall that the discrete-time Fourier transform of $x[n]$ is given by

$$X(\Omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n}. \quad (2.64)$$

From (2.63) and (2.64) we see that the discrete-time Fourier transform is equal to the Z-transform for $z = e^{j\Omega}$, that is,

$$X(z)|_{z=e^{j\Omega}} = \Im\{x[n]\} = X(\Omega). \quad (2.65)$$

Similarly to the Laplace transform, the algebraic expression of the Z-transform of a given function can be equal to the expression corresponding to a different function. Then, Z-transform also demands the specification of the range of values of the complex quantity z for which the transform converges. This range of values is called the *region of convergence* (ROC) of the Z-transform. As an example, consider the Z-transform of $x[n]$ and $y[n]$, as shown below [13, p. 166]:

$$\begin{aligned}
1. \quad x[n] &= a^n u[n], a \text{ real} \Rightarrow X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} = \frac{z}{z-a}, |z| > |a| \\
2. \quad y[n] &= -a^n u[-n-1], a \text{ real} \Rightarrow Y(z) = \sum_{n=-\infty}^{\infty} y[n]z^{-n} = \frac{z}{z-a}, |z| < |a|.
\end{aligned}
\tag{2.66}$$

Observe that, similarly to the Laplace transforms in (2.62), the algebraic expressions of the Z-transform of $x[n]$ and $y[n]$ are equal, but they differ in their ROC.

2.3 Sampling of Deterministic and Random Signals

Sampling is an essential operation for converting a continuous-time signal into a sequence, enabling digital processing of this sequence. In this section we review the main concepts associated to the sampling theory when applied to deterministic and stochastic signals.

2.3.1 Ideal Sampling of Deterministic Signals

A continuous-time (possibly complex) deterministic signal $x(t)$ is considered to be band-limited if the magnitude of its Fourier transform $|X(f)| = 0$ for $|f| > B$, where B is the highest frequency component of $x(t)$. From a practical perspective, it suffices that $X(f)$ is *approximately* zero for $|f| > B$.

Let a sequence of unit impulses and its Fourier transform be written as

$$\sum_{k=-\infty}^{\infty} \delta(t - kT_s) \Leftrightarrow \frac{1}{T_s} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T_s}\right). \tag{2.67}$$

Now let us construct the sampled version of $x(t)$, which we call $x_s(t)$, by multiplying $x(t)$ by the sequence of unit impulses, according to

$$x_s(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT_s). \tag{2.68}$$

Since a multiplication in the time-domain corresponds to a convolution in the frequency-domain, the Fourier transform of $x_s(t)$ will be given by

$$X_s(f) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X\left(f - \frac{k}{T_s}\right). \tag{2.69}$$

Figure 2.6 illustrates the ideal sampling process just described. Note that the spectrum of a sampled signal corresponds to replications of the original spectrum at multiples of the sampling rate $1/T_s$, scaled by $1/T_s$.

The original signal can be completely recovered from its samples if $x_s(t)$ is filtered by an ideal low-pass *reconstruction filter* with gain T_s between $-B$ and $+B$ and zero otherwise, as long as the sampling rate is greater than $2B$ samples per second. This is the so-called *Nyquist sampling criterion*. If the sampling rate is $2B$ samples per second we say that we are sampling at the *Nyquist rate*.

For a sampling rate less than $2B$ samples per second, the spectral replicas in $X_s(f)$ are superimposed, causing a phenomenon called *aliasing*. Aliasing clearly causes distortion and the original signal can not be recovered.

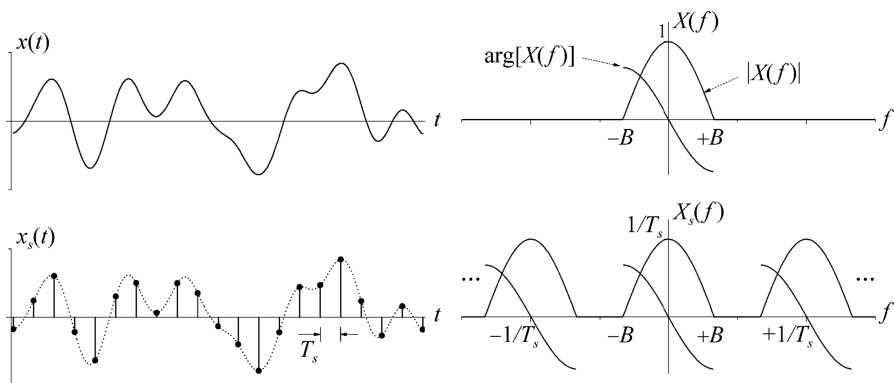


Fig. 2.6 Illustration of the ideal sampling process

The impulse response of the ideal reconstruction filter is given by

$$h(t) = \text{sinc}(2Bt) = \frac{\sin(2\pi Bt)}{2\pi Bt}. \quad (2.70)$$

Then we have a perfectly reconstructed signal

$$x(t) = \sum_{k=-\infty}^{\infty} x\left(\frac{k}{2B}\right) \text{sinc}\left[2B\left(t - \frac{k}{2B}\right)\right], \quad (2.71)$$

where $\{x(k/2B)\}$ are the samples of $x(t)$ taken in integer multiples of $1/(2B)$ seconds. This reconstruction is said to be perfect since

$$x(t) - \sum_{k=-\infty}^{\infty} x\left(\frac{k}{2B}\right) \text{sinc}\left[2B\left(t - \frac{k}{2B}\right)\right] = 0. \quad (2.72)$$

Equation (2.72) shows the reason for the name *ideal interpolating function* usually given to the *sinc* function in the context of sampling. In other words, the samples

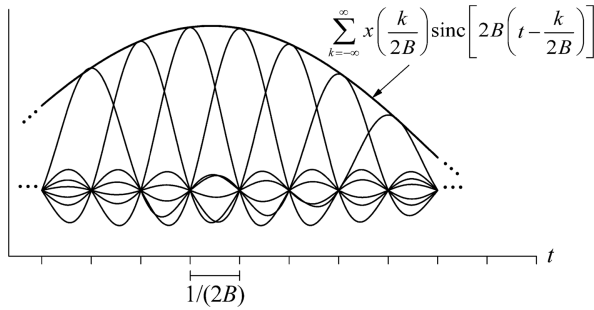


Fig. 2.7 Illustration of the ideal reconstruction of $x(t)$ from its samples

are connected according to the *sinc* functions that are added in-between them. Figure 2.7 illustrates this ideal reconstruction process.

2.3.2 Ideal Sampling of Stochastic Processes

Let $X(t)$ be a band-limited stationary random process, such that $|S_X(f)| = 0$ for $|f| > B$, where B is the highest frequency component of $X(t)$. If this process is sampled at a rate greater than $2B$ samples per second, it can be reconstructed from its samples by using the ideal interpolating function, as shown by

$$Y(t) = \sum_{k=-\infty}^{\infty} X\left(\frac{k}{2B}\right) \text{sinc}\left[2B\left(t - \frac{k}{2B}\right)\right]. \quad (2.73)$$

The reconstruction above is not perfect as in the case of a deterministic signal and the equality between $X(t)$ and $Y(t)$ occurs in the mean square sense, which means that

$$E \left\{ \left| X(t) - \sum_{k=-\infty}^{\infty} X\left(\frac{k}{2B}\right) \text{sinc}\left[2B\left(t - \frac{k}{2B}\right)\right] \right|^2 \right\} = 0. \quad (2.74)$$

2.3.3 Practical Sampling

The ideal sampling process considered previously is not realizable in practice due to the fact that the sequence of Dirac delta impulses is not realizable. Additionally, the ideal reconstruction process is not realizable due to the fact that the ideal interpolating filter is not realizable.

Another problem arises due to the fact that, in practice, most of the signals of interest do not have their spectrum strictly confined in the $(-B, +B)$ range. Then, some residual aliasing will always occur.

Figure 2.8 shows a block diagram of a typical discrete-time processing of a continuous-time signal. The signal $x(t)$ enters an *anti-aliasing filter*, whose objective is to attenuate as much as possible the frequency components of $x(t)$ that can be disregarded. This will reduce the effect of aliasing, also permitting the reduction of the sampling rate.

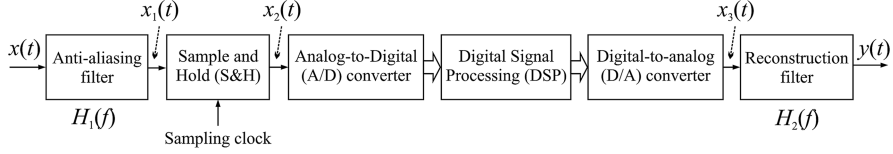


Fig. 2.8 A common digital signal processing of a continuous-time signal

At the output of the anti-aliasing filter, the signal $x_1(t)$ is sampled and held, which means that the value of $x_1(t)$ at the sampling instant is maintained until the next sample instant. This process is also known as a *zeroth-order hold* [35, p. 46].

The signal at the output of the S&H block will look like the one shown in Fig. 2.9. Note that $x_2(t)$ can be written in terms of the ideal sampling as

$$x_2(t) = p(t) * x_1(t) \sum_{k=-\infty}^{\infty} \delta(t - kT_s), \quad (2.75)$$

where $p(t)$ is a rectangular pulse of unit amplitude and duration $T_s < 1/(2B)$ seconds. In the frequency-domain we have

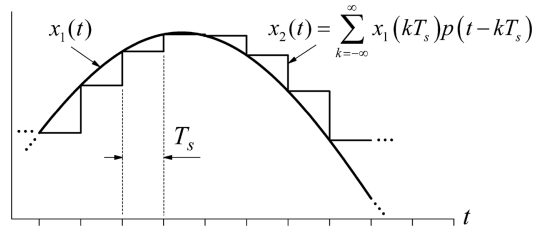


Fig. 2.9 The sample and hold (S&H) process

$$\begin{aligned} X_2(f) &= P(f) \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X_1\left(f - \frac{k}{T_s}\right) \\ &= \left[T_s \frac{\sin(\pi f T_s)}{\pi f T_s} e^{-j\pi f T_s} \right] \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X_1\left(f - \frac{k}{T_s}\right) \\ &= \text{sinc}(f T_s) e^{-j\pi f T_s} \sum_{k=-\infty}^{\infty} X_1\left(f - \frac{k}{T_s}\right). \end{aligned} \quad (2.76)$$

For a better understanding of (2.76), assume that no digital processing is being performed and that no error is produced by the analog-to-digital (A/D) conversion. In this case $x_3(t)$ will be equal to $x_2(t)$. If we additionally want $y(t) = x_1(t)$, the reconstruction filter must compensate for the distortion caused by the multiplication between the replicas of $X_1(f)$ by the $\text{sinc}(fT_s)$ function in (2.76). Then, the signal $x_1(t)$ can be reconstructed from its zeroth-order samples if the signal $x_2(t)$ go through an ideal low-pass filter whose magnitude response is given by

$$|H_2(f)| = \begin{cases} \frac{\pi f T_s}{\sin(\pi f T_s)}, & |f| \leq B \\ 0, & |f| > B. \end{cases} \quad (2.77)$$

This filter is called a $x/\sin(x)$ -type *correction filter*. In practice, the ideal filtering given in (2.77) can be relaxed by adopting a realizable filter having the $x/\sin(x)$ behavior for $|f| \leq B$ and a high attenuation beyond B Hz. The phase response of this filter is not of concern, as long as it is linear for $|f| \leq B$. Figure 2.10 synthesizes the above concepts.

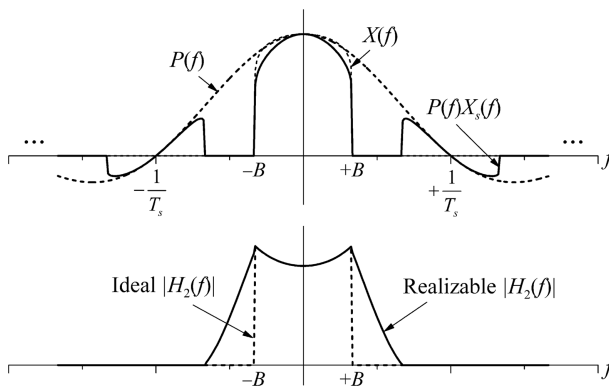


Fig. 2.10 The use of a $x/\sin(x)$ -type correction filter when a zeroth-order sampling is performed

In the upper part of Fig. 2.10, observe the effect of the multiplication between $P(f)$ and $X_s(f)$, caused by the use of the zeroth-order sampling. Observe the distortion in $P(f)X_s(f)$ when compared to the original $X(f)$. To compensate for this distortion, the $x/\sin(x)$ -type reconstruction filter $|H_2(f)|$ is used. In the lower part of Fig. 2.10 it is shown the ideal and a realizable magnitude response of this filter.

In real cases in which some error is introduced by the A/D conversion and some digital signal processing is performed, the $x/\sin(x)$ -type reconstruction filter must be used as long as some digital-to-analog (D/A) conversion is needed.

The error in the A/D conversion is mainly caused by the quantization process, giving rise to the so called quantization error. This topic will be addressed in more detail in the next subsection.

Simulation 2.3 – Sampling



File – CD drive: \Simulations\Signals\Sampling.vsm.

Default simulation settings: Frequency = 20,000 Hz; End = 0.4 second. Sampling rate: 2,000 samples/second.

This experiment aims at illustrating the main concepts related to the sampling theory. A high-bandwidth signal is low-pass filtered so that it can be sampled and produce low aliasing. Most of the power in the filtered signal is concentrated within a bandwidth of 400 Hz.

The filtered signal is sampled in two ways: (1) “ideal” sampling, which approximates the ideal sampling process realized by a sequence of impulses, and (2) zeroth-order sampling. The sampling rate can be configured simultaneously for both sampling processes.

In a situation where the discrete-time processing must provide a continuous-time result, a digital-to-analog conversion must take place. To simulate this process, the zeroth-order samples are low-pass filtered in two ways: (1) using a simple low-pass filter with flat magnitude response and linear phase within the signal bandwidth, and (2) using a low-pass filter with linear phase and a magnitude response with the $x/\sin(x)$ -type equalization. The resultant waveforms are compared against the original continuous-time waveform, using as a criterion the mean-square error between the reconstructed and the original signals.

The simulation also permits the visualization of the frequency response of the reconstruction filters, along with several other plots.

Using the default settings, run the simulation while observing Plot A, just to see the bandwidth limitation imposed by the anti-aliasing filter.

Now open Plot B. The upper graph shows the “ideal” samples of the signal and the lower graph shows the frequency spectrum of these samples. Notice that the ideal sampling generates replicas of the original signal in multiples of the sampling rate, which in this case is 2,000 samples/second. Compare these results with those shown in Plot C, which refers to the sample and hold process. Observe that the frequency spectrum is now modified by the $\sin(x)/x$ shape of $P(f)$, the Fourier transform of the rectangular zeroth-order pulse. Since this pulse has a duration of $T_s = 1/(2,000)$ s, the spectral nulls of $P(f)$ occurs in multiples of 2,000 Hz, as can be observed via Plot C.

As a simple exercise, justify the appearance of negative amplitudes in the frequency spectrum of $P(f)X(f)$ shown in Fig. 2.10 and the absence of such negative amplitudes in the spectrum shown via Plot C.

Plot D shows the reconstructed waveforms along with the original waveform. The outputs of the reconstruction filters were adjusted so that all three signals have the same average power. This allows for a fair comparison among these signals, as shown below.

Observe the mean square error (MSE) between the reconstructed signals and the filtered continuous-time signal. Note that their values are both small, but the $x/\sin(x)$ equalization has produced a better result. The difference between the mean square errors is not higher because of the small slope of $P(f)$ in the bandwidth of $X(f)$. Moreover, the intensity of $X(f)$ in the region of greater distortion (next to its bandwidth edge) is also small. As a result, the overall distortion is small. The closer the sampling rate to the Nyquist rate and the higher the intensity of $X(f)$ in its edged, the greater the distortion caused by the zeroth-order sampling. For a better understanding of these comments, please refer again to Fig. 2.10.

Now open Plot E so that you can compare the frequency responses of the reconstruction filters. Note that they have approximately the same cutoff frequency, but slightly different stop-band attenuations. Magnify their pass bands so that you can observe the flat response of one filter and the $x/\sin(x)$ equalization curve of the other. The difference in the pass bands can be more easily observed if you uncheck the log scale for the y-axis of the plot. You will have to rescale the plot after that in order to have a good visualization.

To see the effect of the sampling rate, change its value to 1,000, then 500 and 250 samples/second and observe plots B, C and D. Note that above 1,000 samples/second aliasing is practically zero. Note also that the MSE between the reconstructed and the original signals becomes progressively higher as the sampling frequency becomes smaller. For 250 samples/second it attains its maximum value of 1, which is equivalent to the comparison of the original signal with 0.

Explore inside the individual blocks. Try to understand how they were implemented. Create and investigate for yourself new situations and configurations of the simulation parameters and try to reach your own conclusions.

As a complement, in what follows we present some short additional comments related to deployments of the sampling theory.

2.3.3.1 Aliasing as a Desired Effect

We have seen that aliasing is an undesired effect, since it distorts the signal reconstructed from its samples. However, there are situations in which aliasing is desired and, in fact, it is the easiest way of solving some specific problems.

When a signal is under-sampled ($1/T_s \leq 2B$), some of its frequency components are aliased to lower frequencies and this fact can be explored favorably. As an example, if a 2 MHz cosine signal is sampled at 3 Mega-samples/second, a frequency component in 1 MHz will appear due to aliasing. This is the principle behind the

sampling oscilloscopes, permitting that the equipment extends its capabilities of showing periodic signals with high frequency components.

2.3.3.2 Sampling of Discrete-Time Signals

Discrete-time signals can also be sampled through methods called *up-sampling* (or interpolation) and *down-sampling* (or decimation). Along with the variation of the sampling rate of a continuous-time signal, up-sampling and down-sampling represent useful tools for controlling the time and frequency characteristics of a signal. Applications of up-sampling and down-sampling include the relaxation in the design of filters operating in the discrete-time domain, the so-called *digital filters*.

For more information on the applications of a controlled aliasing and on the up-sampling and down-sampling processes, the interested reader is encouraged to consult classical references on signals and systems and discrete-time signal processing, for example [11], [23] and [24].

2.3.3.3 Sampling of Passband Signals

Finally, it is worth mentioning that the sampling rate of a passband signal does not need to be twice higher than its highest frequency component. The Nyquist sampling theorem applied to passband signals states that if a real passband signal has nonzero frequency content in a range of $B = f_2 - f_1$ Hz, the signal can be reconstructed from its samples if the sampling rate is greater than $2B$ samples per second. The restriction for the lower limit of $2B$ samples per second is that f_2 must be an integer multiple of B [27, pp. 742–746].

The reconstruction process of a sampled passband signal is a little bit more intricate than that used for baseband signals, as indicated in [6, p. 245] and references therein.

2.3.4 Analog-to-Digital Conversion

Digital signal processing techniques are used today in the majority of functions performed in digital communication system hardware. Since the concept of *software-defined radio* (SDR)⁴ was born, analog-to-digital (A/D) and digital-to-analog (D/A) conversions have been moved towards the receiving and transmitting antennas, respectively, thus becoming more critical tasks due to the high frequencies involved.

Because of its importance in the discrete-time processing of continuous-time signals, A/D conversion will be studied in more detail in this subsection. Digital-to-analog (D/A) conversion is also an important part of this processing. Nevertheless, the most critical errors are caused by the A/D conversion.

⁴ For an overview about the SDR concept, see [38]. For other aspects of the SDR, see [3].

A/D conversion is realized in two steps: first the levels of the input samples are approximated to a finite number of levels in a process called *quantization*. Then, each of the N quantized values is converted into a word of $\log_2(N)$ bits. If the quantization levels are equally-spaced we have a *uniform quantization*. If they are not equally-spaced, we have a *non-uniform quantization*.

The quantization process produces an error between the quantized and the non-quantized signals, which is called *quantization error*.

2.3.4.1 Uniform Quantization

Figure 2.11 illustrates a uniform quantization process. In this illustration the error is limited in the range $[-q/2, +q/2]$, as long as the input signal does not exceed the A/D conversion dynamic range of $2X_m$. When the input signal exceeds $2X_m$ the quantization error increases progressively because the quantized signal remains unchanged. This effect is called *clipping* and it is also illustrated in Fig. 2.11.

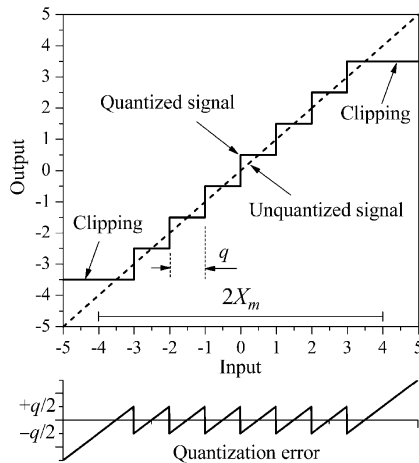


Fig. 2.11 The uniform quantization process and the quantization error

The quantization error can be reduced by increasing the *resolution* of the A/D converter. This resolution can be specified in terms of the quantization step-size q or in terms of the number of bits that represents each sample value.

For the majority of signals having complex excursions, like voice signals and other random signals, the quantization error will be uniformly-distributed in the range $[-q/2, +q/2]$. Since it has zero mean, its variance (which in this case corresponds to its average power) is computed according to

$$\sigma_e^2 = \int_{-q/2}^{+q/2} z^2 \frac{1}{q} dz = \frac{q^2}{12}. \quad (2.78)$$

The relation between the quantization step-size and the dynamic range of the A/D converter is given by

$$q = \frac{2X_m}{2^b}, \quad (2.79)$$

where b is the number of bits (or resolution) of the A/D converter. With the above results at hand we are able to define the *signal-to-quantization noise ratio* (SNR_Q), as shown by:

$$\begin{aligned} \text{SNR}_Q &= 10 \log \left(\frac{\sigma_X^2}{\sigma_e^2} \right) = 10 \log \left(\frac{12\sigma_X^2}{q^2} \right) = 10 \log \left[3 \left(\frac{\sigma_X}{X_m} \right)^2 2^{2b} \right] \\ &= 6.021b + 4.771 - 20 \log \left(\frac{X_m}{\sigma_X} \right) \text{ dB}. \end{aligned} \quad (2.80)$$

For a signal to be represented by the majority of the quantization levels, its excursion must fit within the dynamic range of the A/D converter. In this case no clipping will occur. However, if a signal has a large peak-to-average value, high peaks will be relatively rare. In this situation, permitting some clipping might be a better choice. Let us investigate this subject more carefully: the *peak-to-rms* ratio or *crest factor* of a waveform is the ratio between the peak amplitude of the waveform and its root mean square value:

$$C_F = \frac{|X|_{\max}}{\sigma_X}. \quad (2.81)$$

Without loss of generality, we adopt the peak of the waveform as being X_m , such that its *rms* value is modified to keep the crest factor unchanged. In this case the crest factor can be written as

$$C_F = \frac{X_m}{\sigma_X}. \quad (2.82)$$

With this result in (2.80), the signal-to-quantization noise ratio becomes

$$\text{SNR}_Q = 6.021b + 4.771 - 20 \log(C_F) \text{ dB}. \quad (2.83)$$

The signal-to-quantization noise ratio does depend on the resolution of the A/D converter, since it is improved by 6 dB for each bit added to the word length. Nevertheless, the SNR_Q also depends on the crest factor of the signal to be quantized. If C_F is high, the SNR_Q and the occurrence of clipping decrease. If C_F is low, the SNR_Q and the occurrence of clipping increase.

Some typical values of $-20 \log(C_F)$ are: -3 dB for a sinusoidal signal; -15 dB for a telephony voice signal and -12 dB for a typical music signal.

The dependence of the SNR_Q on the crest factor of the signal to be quantized is translated by (2.80) into a long-term average. In other words, (2.80) does not

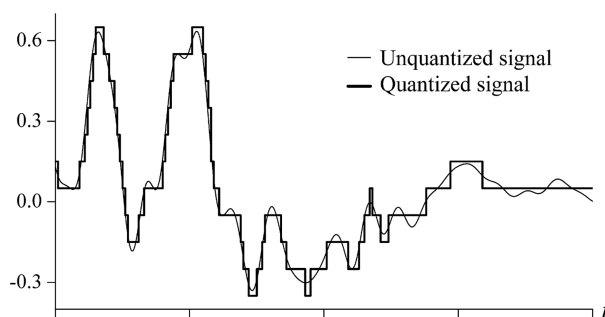


Fig. 2.12 The effect of uniform quantization in the short-term SNR_Q

reveal the short-term influence of the quantization error. To illustrate this, suppose that a signal with high crest factor is to be uniformly quantized. Due to the high crest factor, low level amplitudes are frequent and, when this occurs, the short-term SNR_Q decreases. This happens due to the fact that uniformly-spaced quantization levels are not able to represent with enough resolution the small variations in the signal amplitude, as illustrated in Fig. 2.12. A solution to this problem is the use of non-uniform quantization, as shown in what follows.

2.3.4.2 Non-uniform Quantization

With non-uniform quantization, the region of lower amplitudes is quantized with higher granularity. This granularity decreases as the signal amplitude increases. In other words, the quantization step-size is small in the region of lower amplitudes and increases in higher amplitudes. As a consequence, short-term (and also long-term) SNR_Q can be kept approximately constant.

Figure 2.13 shows the same unquantized signal shown in Fig. 2.12, but now a non-uniform quantization process is illustrated. Note that lower signal amplitudes are quantized with a higher granularity and that higher amplitudes are quantized with far apart quantization levels.

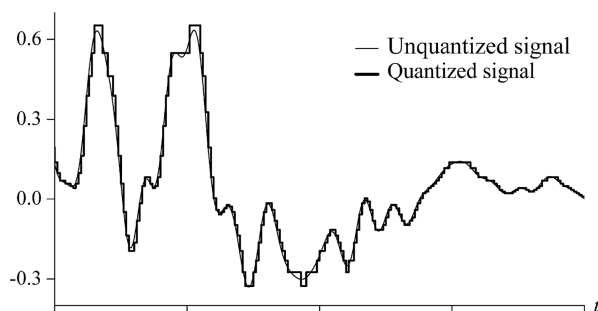


Fig. 2.13 The effect of non-uniform quantization in the short-term SNR_Q

One way of achieving the effect of a non-uniform quantization is to compress the input signal using some logarithmic rule, and then apply the compressed signal to a uniform quantizer. The resultant continuous-time output signal is then expanded by using the complementary behavior of the compression curve. The combination of the compression and expansion characteristics is usually referred to as *companding*.

Typical compression rules are the μ -law and the A -law. The μ -law is used mainly in the USA and the A -law is used mainly in Europe [34, pp. 621, 625]. In what follows we consider only the A -law as a case study.

The relation between the input (x) and the output (y) of an A -law compression curve is given by

$$y = \begin{cases} y_{\max} \frac{A(|x|/x_{\max})}{1 + \ln A} \operatorname{sgn} x, & 0 < |x|/x_{\max} < 1/A \\ y_{\max} \frac{1 + \ln[A(|x|/x_{\max})]}{1 + \ln A} \operatorname{sgn} x, & 1/A < |x|/x_{\max} < 1, \end{cases} \quad (2.84)$$

where $\operatorname{sgn}(x)$ is the *sign* function or *signum* function, defined by

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases} \quad (2.85)$$

The parameter A in (2.84) governs the compression ratio, as illustrated in Fig. 2.14. In practice, the compression curve is implemented as linear piecewise chords, facilitating the realization of the A -law compression [34, p. 625].

The signal-to-quantization noise ratio will become independent of the crest factor and the level of the input signal, as shown by [6, p. 150]:

$$\operatorname{SNR}_Q = 6.021b + 4.771 - 20 \log(1 + \ln A) \text{ dB}. \quad (2.86)$$

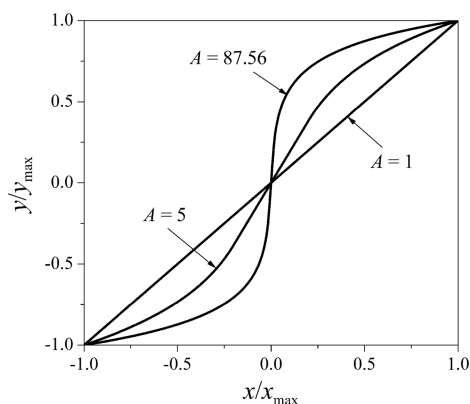


Fig. 2.14 Compression characteristics of the A -law

A typical value of A is 87.56. With this value and an 8-bit A/D converter, the average SNR_Q obtained from (2.86) is approximately 38 dB.

The next simulation aims at exploring more details about the uniform and non-uniform quantization processes. For a comprehensive and very didactical treatment on the subject, see Chap. 11 of [34].

Simulation 2.4 – Uniform and Non-uniform Quantization



File – CD drive: \Simulations\Signals\Quantization.vsm.

Default simulation settings: Frequency = 200 Hz; End = 20 seconds.
Half of the dynamic range: 3.161 volts. Gain: 1.3.

This experiment aims at confirming some theoretical results and complementing the concepts related to the uniform and non-uniform quantization methods.

Open the simulation file indicated in the header and notice that the experiment is composed by two independent parts:

- In the upper part, a ramp signal is applied to uniform and non-uniform quantizers. The non-uniform quantizer is an A-law quantizer with $A = 87.56$. The output signal and the quantization noise produced by these quantizers are analyzed via plots A and B.
- In the lower part of the experiment a sample-function of a voice signal is quantized by a 6-bit uniform quantizer and by a 6-bit non-uniform quantizer. The A-law quantizer with $A = 87.56$ is also used in this part of the experiment. The dynamic range of these quantizers can be configured via the “half of the dynamic range” block. The signals and quantization errors are analyzed through plots C, D, E and F. The voice signal can have its amplitude configured in the “gain” block, and the signal crest factor can be computed through the estimates of its peak and *rms* values. The corresponding signal-to-quantization noise ratios are also estimated.

Using the default configurations, run the simulation while observing Plot A. Note that it reproduces the behavior shown in Fig. 2.11. Here the input signal is a ramp going from -10 to $+10$ volts. The quantizer has 8 levels, so it is part of a 3-bit A/D converter. The dynamic range is $2X_m = 8$ V, from -4 to $+4$ V and the quantization step-size is, according to (2.79), $q = 8/8 = 1$ volt. Note that beyond the limits imposed by the dynamic range the quantized signal is clipped to the extreme quantization levels of -3.5 or $+3.5$ V. Observe also that the quantization error is limited to $\pm q/2$, except when the input signal exceeds the dynamic range.

Now, while observing Plot B, run the simulation again. Compare the plots with those in Plot A and note that the quantization step-sizes are not equally spaced anymore. In fact, they are smaller for small input signal amplitudes and larger for large input signal amplitudes. Recall that this behavior is characteristic of a non-uniform quantization. Observe also that the quantization error is smaller for small input signals and larger for large input signals.

Now let us move our attention to the lower part of the experiment. Run the simulation while looking at Plot D. Observe the time plot and the histogram of the quantization error. Note that we are using the default gain of 1.3 and, in this case, the peak of the voice signal is $X_m = 3.161$ volts. Then, according to (2.79), $q = 2 \times 3.161/(2^6) \cong 0.1$ volts. Since we are analyzing the uniform quantizer, these plots show to us that, in fact, the quantization error is uniformly-distributed in the range $[-q/2, +q/2]$. Note also that the voice signal seems to have a Laplace distribution. From Chap. 1 we recall that a Laplace PDF and its root mean square value are

$$\begin{aligned} f_X(x) &= \frac{1}{2r} e^{-|x|/r}, \quad r > 0 \\ \sigma_X &= r\sqrt{2}. \end{aligned} \quad (2.87)$$

By plotting a Laplace PDF with scale parameter $r \cong 0.56$ and mean $\mu = 0$ over the voice signal histogram, as shown in Fig. 2.15, we observe a good agreement between them. From (2.87), the theoretical *rms* value of the voice signal is $\sigma_X \cong 0.792$ volts. This result agrees with the *rms* value estimated in the simulation.

Since the peak of the voice signal is $X_m = 3.161$ volts, the crest factor is $C_F \cong 3.16/0.79 = 4$. With this result in (2.83) we obtain

$$\text{SNR}_Q = 6.021 \times 6 + 4.771 - 20 \log(4) \cong 28.94 \text{ dB}. \quad (2.88)$$

Verify that this result is approximately equal to the long-term average signal-to-quantization noise ratio estimated in the simulation.

Now, run the simulation while observing Plot C. With a right-click over each plot, configure them using the following limits for the X and Y axis:

- Y upper bound: 0.7
- Y lower bound: -0.4
- X upper bound: 18.14
- X lower bound: 17.14

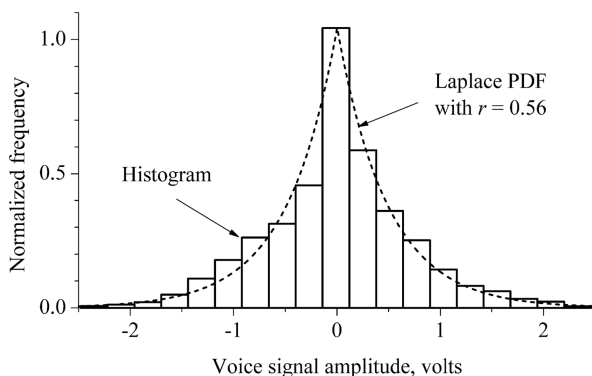


Fig. 2.15 Normalized histogram of the voice signal and an over-plotted Laplace PDF

After this configuration is done, click the OK button. You will see exactly the same waveforms shown in Fig. 2.12. Of course, the comments related to that figure also apply here.

We now move to the analysis of the non-uniform quantization. Run the simulation while observing Plot F. First, notice that the quantization error is not uniformly-distributed anymore. Coincidentally, it seems to be Laplace-distributed. Secondly, note that the quantization error is smaller when the voice signal has lower amplitudes, a characteristic that is intrinsic to the non-uniform quantization.

Open Plot E and configure the graphs according to the axis limits given above. Click the OK button and note that the waveforms are exactly the same as the ones shown in Fig. 2.13. Again, the comments concerning that figure also apply here.

The theoretical signal-to-quantization noise ratio considering the A-law compression rule is given by (2.86). Using the simulation parameters we obtain

$$\text{SNR}_Q = 6.021 \times 6 + 4.771 - 20 \log[1 + \ln(87.56)] \cong 26.13 \text{ dB}. \quad (2.89)$$

Observe that this result is also approximately equal to the long-term average signal-to-quantization noise estimated in the simulation. Note that the SNR_Q for the uniform quantizer is slightly higher than the SNR_Q for the non-uniform quantizer. This seems to be unconditionally good, but it is not. The lower long-term SNR_Q produced by the non-uniform quantizer is compensated by the fact that it is almost independent of the input signal amplitude. This is an effect that is desired mainly in cases where the crest factor is high, since in these cases the probability of occurrence of lower amplitude levels increases.

Now open Plot G. This plot is recording the short-term variation of the signal-to-quantization noise ratio for both the uniform and the non-uniform quantizers. This short-term SNR_Q is being estimated through a sliding window average. In this form of estimation the signal power and the noise power are measured during a shorter time compared to the entire simulation time. Specifically, we are using a window with 100 simulation steps, which corresponds to the following time interval: $(\text{number of steps}) \times (1/\text{simulation frequency}) = 100 \times (1/200) = 0.5$ seconds. This means that a window of 0.5 seconds slides over the signal and the quantization noise while their powers are estimated.

As previously mentioned, non-uniform quantization produces a short-term SNR_Q that is approximately constant, while uniform quantization produces low values of SNR_Q when the input signal has lower amplitudes and high values of SNR_Q when the input signal has higher amplitudes. This can be clearly observed through Plot G, confirming the theoretical expected behavior.

Finally, change the amplitude of the voice signal via the “gain” block. If you modify the value of the dynamic range accordingly, everything happens as before, except by the fact that the quantization step-size is changed to keep the number of quantization levels in 64, which corresponds to a 6-bit A/D converter. If the dynamic range is set with a value lower than the peak amplitude of the voice signal, clipping will occur, leading to a severe degradation in the signal-to-quantization noise ratios for both the uniform and the non-uniform quantizers. This happens due to the fact

that the quantization error increases dramatically when clipping occurs. Implement these tests and analyze the resultant SNR_Q .

Explore inside the individual blocks. Try to understand how they were implemented. Create and investigate for yourself new situations and configurations of the simulation parameters and try to reach your own conclusions.

2.3.4.3 Other A/D Conversion Techniques

We can say that the approach for the digital-to-analog conversion adopted as reference for the study in this book is a classical or conventional approach. There are several others A/D conversion techniques in which the quantization error always exists, but it manifests differently from that we have presented before. Examples of these alternative methods are [20]: the *parallel* or *flash*, the *successive approximation*, the *pipeline* and the *cyclic* A/D conversions. One of the most used techniques, mainly in telecommunications and audio, is the oversampled *single-bit* (or 1-bit) A/D converter. In this converter a higher sampling rate is traded for a lower number of bits. Most of the single-bit A/D converters are based on the delta modulation [6] and, for this reason, they are also known as *delta-sigma* or *sigma-delta* converters [2, 20, 25, 30, p. 2227].

2.4 Linear Systems

A system is a process that transforms the input, called *excitation*, into an output, called *response*, through some mapping function.

In this section we present an important class of systems called *linear systems*. They play a major role in the modeling and analysis of communication systems and the related concepts will be extensively used throughout the rest of the book.

We start by characterizing linear systems in the time-domain and later we move to the frequency-domain characterization. We end the section by addressing the main properties of linear systems.

2.4.1 Time-Domain Characterization of Linear Systems

Consider the system model depicted in Fig. 2.16, which represents a continuous-time or a discrete-time system, depending on the nature of the inputs and outputs and of the system itself. It is assumed that this system is time-invariant, that is, if $y(t)$ is the system response to an input $x(t)$, the response to a delayed input $x(t - \tau)$ will be the corresponding delayed version of $y(t)$, that is, $y(t - \tau)$.

In the context of communication theory, we are particularly interested in a class of systems called *linear systems*. These systems satisfy the properties of *superposition* and *scaling*, as shown in the sequel.

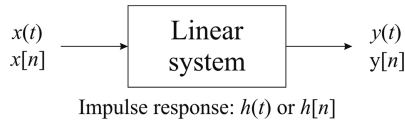


Fig. 2.16 Linear system model

Let $x(t) = x_1(t) + x_2(t)$ and let $y_1(t)$ be the system response when $x_1(t)$ is the excitation and $y_2(t)$ be the system response when $x_2(t)$ is the excitation. Also let $y(t) = T[x(t)]$, where $T(\cdot)$ is a function that maps the system input into the system output. The superposition and scaling properties applied to this scenario can be written respectively as follows:

$$\begin{aligned} 1. \quad & T[x_1(t) + x_2(t)] = y_1(t) + y_2(t) \\ 2. \quad & T[\alpha x(t)] = \alpha y(t). \end{aligned} \tag{2.90}$$

We can interpret (2.90) in the following way: the superposition property is associated to the fact that, if $x(t) = x_1(t) + x_2(t)$, the response $y(t)$ can be determined by applying $x_1(t)$ to the input, obtaining the output $y_1(t)$ and then applying $x_2(t)$ to the input, obtaining the output $y_2(t)$. The total output $y(t)$ will be given by the sum of the individual responses $y_1(t)$ and $y_2(t)$. The scaling property states that if we multiply the input signal by a given constant, the output is multiplied by the same amount. These properties can be generalized according to:

$$\text{If } x(t) = \sum_k \alpha_k x_k(t), \text{ then } y(t) = \sum_k \alpha_k y_k(t). \tag{2.91}$$

A system that does not satisfy (2.91) is said to be a *nonlinear system*. The mathematical treatment of nonlinear systems is usually more complicated than for the case of linear systems. We devote our attention to linear systems due to their frequent use to model communication systems.

The output signal $y(t)$ or $y[n]$ in the linear system of Fig. 2.16 is mapped from the input signal $x(t)$ or $x[n]$ through the *impulse response* of the system, $h(t)$ or $h[n]$. The term impulse response means that if we apply a Dirac delta function to the input of a continuous-time linear system and a Kronecker delta function to the input of a discrete-time linear system, the outputs are $h(t)$ and $h[n]$, respectively. Then, from (2.10) we can write

$$x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau. \tag{2.92}$$

Applying the mapping operator we obtain

$$y(t) = T[x(t)] = T \left[\int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau \right]. \tag{2.93}$$

Due to the linearity property we can rewrite (2.93) as

$$y(t) = \int_{-\infty}^{\infty} x(\tau)T[\delta(t - \tau)]d\tau. \quad (2.94)$$

But if the impulse response $h(t)$ does not vary with time and, by definition $T[\delta(t)] = h(t)$, then $T[\delta(t - \tau)] = h(t - \tau)$. With this result in (2.94) we finally obtain

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau. \quad (2.95)$$

This operation is the *convolution integral* between the input signal $x(t)$ and the system impulse response $h(t)$. It is represented by

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau. \quad (2.96)$$

Similarly, for a discrete-time system we have

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} h[k]x[n - k] = \sum_{k=-\infty}^{\infty} x[k]h[n - k]. \quad (2.97)$$

In (2.97) we have the *convolution sum* between the discrete-time input signal $x[n]$ and the discrete-time system impulse response $h[n]$. The operations in (2.96) and (2.97) describe the system's input/output relationship in the time-domain. Let us move to the frequency-domain analysis.

2.4.2 Frequency-Domain Characterization of Linear Systems

Now we focus our attention to the frequency-domain description of linear systems, which is built on the Fourier analysis theory. First we define the frequency response of a linear system as the Fourier transform of its impulse response, as shown below for a continuous-time and for a discrete-time system, respectively:

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}dt \quad \text{and} \\ H(\Omega) &= \sum_{n=-\infty}^{\infty} h[n]e^{-j\Omega n}. \end{aligned} \quad (2.98)$$

We must recall that, for continuous-time systems, the *angular frequency* (in radians per second) is $\omega = 2\pi f$, where f is measured in hertz. As we shall see later,

for discrete-time systems the angular frequency (in radians) is $\Omega = 2\pi f T_s$, where T_s is the interval between discrete-time samples.

In general, $H(f)$ and $H(\Omega)$ are complex functions and the frequency responses in (2.98) can be written as

$$\begin{aligned} H(f) &= |H(f)|e^{j\Theta(f)} \quad \text{and} \\ H(\Omega) &= |H(\Omega)|e^{j\Theta(\Omega)}, \end{aligned} \tag{2.99}$$

where $|H(f)|$ and $|H(\Omega)|$ are the *magnitude of the frequency responses*, or simply the *magnitude responses*, and $\Theta(f)$ and $\Theta(\Omega)$ are the *phase of the frequency responses*, or simply the *phase responses* of the system.

The convolution integral and the convolution sum were used to determine the output of a linear system in the time-domain. In the frequency-domain, these operations are converted into multiplications, as shown by:

$$\begin{aligned} Y(f) &= X(f)H(f) \quad \text{and} \\ Y(\Omega) &= X(\Omega)H(\Omega). \end{aligned} \tag{2.100}$$

Sometimes it is easier to perform a frequency-domain multiplication than to perform a time-domain convolution. The time-domain functions $y(t)$ and $y[n]$ can be determined from the inverse Fourier transforms of $Y(f)$ and $Y(\Omega)$.

Since the frequency response of a discrete-time system, $H(\Omega)$, is determined by the discrete-time Fourier transform of $h[n]$ via the second expression in (2.98), it is periodic with period 2π , that is,

$$H(\Omega + 2\pi) = H(\Omega). \tag{2.101}$$

Then, we must consider the values of Ω (in radians) only over $[0, 2\pi)$ or $[-\pi, \pi)$ in $H(\Omega)$. In the continuous-time case, the frequency f (in hertz) can be considered in the infinite range $(-\infty, +\infty)$. Later on in this section we shall consider the mapping between these two frequency measures in a way that a discrete-time system can be used to model a continuous-time system.

2.4.2.1 Distortion-Free Linear System

The condition for distortion-free transmission through a linear system can be written as

$$y(t) = x(t) * h(t) = \kappa x(t - \tau), \tag{2.102}$$

where $y(t)$ is the signal at the system output, $x(t)$ is the input signal, $h(t)$ is the system impulse response and $\kappa \neq 0$ is an arbitrary constant. In words, if the output of the system is only a scaled and delayed version of the input waveform, no distortion has occurred. Writing (2.102) in the frequency domain we obtain

$$Y(f) = X(f)H(f) = \kappa X(f) \exp(-j2\pi f \tau), \quad (2.103)$$

from where we can determine the frequency response of the system in order to cause no distortion to the signal, as shown by:

$$\begin{aligned} H(f) &= \frac{Y(f)}{X(f)} \\ &= \frac{\kappa X(f) \exp(-j2\pi f \tau)}{X(f)} \\ &= \kappa \exp(-j2\pi f \tau), \end{aligned} \quad (2.104)$$

from where we can write

$$\begin{aligned} |H(f)| &= \kappa \quad \text{and} \\ \Theta(f) &= -2\pi f \tau. \end{aligned} \quad (2.105)$$

From (2.105) we conclude that, for distortion-free transmission through a linear system, the magnitude of its frequency response must be constant and its phase response must be a linear function of f . The constant frequency response means that all frequency components of the signal are affected by the same gain. The linear behavior of the phase response means that each frequency component of the signal is affected by the same delay from the system input to the system output.

2.4.2.2 Group Delay

The phase response of a linear system can reveal another important parameter for system characterization: the *group delay*. For a continuous-time system, it is defined as the negative of the derivative of the phase response with respect to the ordinary frequency f , or to the angular frequency $\omega = 2\pi f$:

$$\tau(f) = -\frac{d}{df} \Theta(f). \quad (2.106)$$

As the name suggests, the group delay is the delay experienced by a small group of frequencies around a given value of f . Note that if the phase response is linear, the group delay is constant, that is, the delay imposed by the system is the same for all frequencies. We then conclude that a non-constant group delay is characteristic of a system that causes distortion.

The group delay is sometimes referred to as the *envelope delay*. However, this term seems to fit better in a distortionless scenario. For example, suppose that the signal $x(t) = a(t) \cos(2\pi f_c t)$ is applied to a distortion-free linear system, where $a(t)$ is a modulating signal that imposes envelope variations in the carrier signal $\cos(2\pi f_c t)$. With $x(t)$ in (2.102) we obtain

$$\begin{aligned}
 y(t) &= \kappa x(t - \tau) = \kappa a(t - \tau) \cos(2\pi f_c t - 2\pi f_c \tau) \\
 &= \kappa a(t - \tau) \cos(2\pi f_c t + \theta),
 \end{aligned}
 \tag{2.107}$$

from where we clearly see that the *envelope* of the signal has been delayed by τ from the input to the output of the linear system. Additionally, this delay has caused the phase of the carrier to be changed in $\theta = -2\pi f_c \tau$ radians, which corresponds to the value of the system phase response for $f = f_c$.

2.4.3 Classifications and Properties of Linear Systems

Linear systems have many interesting and useful properties that reveal their multiple facets and expand the spectrum of their applications. In what follows we discuss some of these properties, as well as the main classifications of such systems.

We concentrate on the continuous-time linear systems. The concepts apply to the discrete-time systems, with the adequate interpretation and change in the mathematical notation. Sometimes the nomenclature also changes slightly to distinguish a discrete-time from a continuous-time case.

2.4.3.1 Linear Time-Invariant (LTI) Continuous-Time Systems⁵

A linear time-invariant (LTI) system has an impulse response that does not vary with time. For an LTI system, if $y(t)$ is the system response to an input $x(t)$, the response to a delayed input $x(t - \tau)$ will be the corresponding delayed version of $y(t)$, that is, $y(t - \tau)$. The theory developed so far has assumed an LTI system.

2.4.3.2 Linear Time-Variant (LTV) Continuous-Time Systems

As the name indicates, a linear time-variant system has an impulse response that varies with time. The output of this system can be determined by the convolution

$$y(t) = \int_{-\infty}^{\infty} h(t, \tau) x(t - \tau) d\tau. \tag{2.108}$$

Physically, the time-variant impulse response $h(t, \tau)$ can be interpreted as the response of the system at time τ due to a Dirac delta function that excited the system at the time $t - \tau$.

As a consequence of its time-varying nature, an LTV system has also a time-varying frequency response which is given by

$$H(f, t) = \int_{-\infty}^{\infty} h(t, \tau) e^{-j2\pi f \tau} d\tau. \tag{2.109}$$

⁵The term *shift-invariant* is the discrete-time counterpart of the *time-invariant* term used to characterize continuous-time systems.

An LTV system characterizes an important class of communication channel: the mobile wireless channel. As we shall see in Chap. 3, the time-variant nature of the mobile wireless channel and its consequences represent one of the main problems to be overcome in a mobile communication system design.

2.4.3.3 Causal Continuous-Time LTI Systems

A causal linear system produces no response before the excitation is applied, which means that, for an LTI system to be causal it must satisfy

$$h(t) = 0, \quad t < 0. \quad (2.110)$$

2.4.3.4 Stable Continuous-Time LTI Systems

An LTI system is said to be stable if a finite maximum absolute value of the input $x(t)$ produces a finite maximum absolute value of the output $y(t)$. In other words, the system output does not diverge for limited inputs. Mathematically, if we have a bounded-input such that

$$|x(t)| < V_1 < \infty, \quad \forall t, \quad (2.111)$$

a stable LTI system produces a bounded-output such that

$$|y(t)| < V_2 < \infty, \quad \forall t, \quad (2.112)$$

where V_1 and V_2 are constants. A system that satisfies (2.111) and (2.112) is said to be BIBO (bounded-input, bounded-output) stable.

A necessary and sufficient condition for a system to be BIBO stable is that the absolute value of its impulse response must be integrable, that is,

$$\int_{-\infty}^{\infty} |h(t)| dt < \infty. \quad (2.113)$$

In the frequency-domain, the necessary and sufficient condition for a system to be BIBO stable is that the region of convergence (ROC) must contain the imaginary axis (see Sect. 2.2.5).

It is worth mentioning that, even if an LTI system does not satisfy (2.113), it may not go into an unstable situation, as long as the input condition that leads to the instability does not manifest. In other words, a system does not enter an unstable condition if the input that yields this condition is not present.

2.4.3.5 Eigenfunctions of Continuous-Time LTI Systems

An *eigenfunction* is a function that, when applied to the input of an LTI system produces the same function as the response, maybe scaled by some amount. Mathematically, a function $x(t)$ satisfying

$$T[x(t)] = \lambda x(t) \quad (2.114)$$

is said to be an eigenfunction of the LTI system. The constant λ is called the *eigenvalue* corresponding to the eigenfunction $x(t)$.

An important eigenfunction is the *complex exponential* $x(t) = \exp(j2\pi f_0 t)$. The response of an LTI system to this function is determined from (2.96) as follows:

$$y(t) = \int_{-\infty}^{\infty} h(\tau) e^{j2\pi f_0(t-\tau)} d\tau = e^{j2\pi f_0 t} \int_{-\infty}^{\infty} h(\tau) e^{-j2\pi f_0 \tau} d\tau. \quad (2.115)$$

Provided that the last integral in (2.115) converges, it represents the system frequency response $H(f)$ for $f = f_0$, which is a constant. Then we can write

$$y(t) = H(f_0) e^{j2\pi f_0 t} = \lambda x(t). \quad (2.116)$$

In words, if we apply a complex exponential to the input of an LTI system, the output is a complex exponential, with the magnitude affected by $|H(f_0)|$ and with the phase affected by $\arg[H(f_0)]$.

This important result can be interpreted as follows: first, recall that a cosine or a sine function can be written in terms of complex exponentials through the Euler relations. Then, if we apply a cosine function to the input of an LTI system, the output is a cosine function, maybe with modified amplitude and phase. Since the cosine and sine functions can be considered as elementary functions of a signal, we conclude that an LTI system does not produce frequency components in addition to the frequency components of the input signal, which is an intuitively satisfying interpretation. Secondly, the response of an LTI system to a signal that can be resolved into a summation of complex exponentials can be determined by using the eigenfunction concept and the superposition property of a linear system.

2.4.3.6 LTI Systems With and Without Memory

An LTI system is said to have memory if its output at a given time instant depends on present and past values of the input. An LTI system is *memoryless* if its output in a given time instant depends only on the input applied at that instant. The memoryless condition for a continuous-time and a discrete-time LTI system implies respectively that

$$\begin{aligned} h(t) &= A\delta(t) \quad \text{and} \\ h[n] &= A\delta[n], \end{aligned} \quad (2.117)$$

where $A = h(0)$ or $A = h[0]$ is a constant.

2.4.3.7 Frequency Response for Systems Characterized by Linear Differential Equations

The input–output relationship of some LTI systems can be characterized by linear constant coefficient differential equations in the form [24, p. 330]:

$$\sum_{k=0}^N a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^M b_k \frac{d^k x(t)}{dt^k}. \quad (2.118)$$

For these systems, the frequency response can be determined by

$$H(f) = \frac{Y(f)}{X(f)} = \frac{\sum_{k=0}^M b_k (j2\pi f)^k}{\sum_{k=0}^N a_k (j2\pi f)^k}. \quad (2.119)$$

If the impulse response $h(t)$ of the system has to be determined, a typical rule is to apply a partial-fraction expansion to (2.119) so that the inverse Fourier transform can be found in a simpler way [24, p. 332]. In fact, partial-fraction expansion is very useful for inverting Fourier, Laplace and Z-transforms in the cases where the transform can be written in terms of rational functions, i.e. in terms of the ratio between two polynomials. For an applied treatment of the partial-fraction expansion in the characterization of linear systems, refer to [11, p. 671; 13, p. 120; 17, p. 24].

2.4.3.8 Frequency Response for Systems Characterized by Linear Difference Equations

Similar to the case of continuous-time LTI systems, the input–output relationship of some LTI discrete-time systems can be characterized by linear constant coefficient difference equations in the form [24, p. 397]:

$$\sum_{k=0}^N a_k y[n-k] = \sum_{k=0}^M b_k x[n-k]. \quad (2.120)$$

For these systems, the frequency response can be determined by

$$H(\Omega) = \frac{Y(\Omega)}{X(\Omega)} = \frac{\sum_{k=0}^M b_k e^{-jk\Omega}}{\sum_{k=0}^N a_k e^{-jk\Omega}}. \quad (2.121)$$

Again, if the impulse response $h[n]$ of the system has to be determined, a typical rule is to apply a partial-fraction expansion to (2.121) so that the inverse Fourier transform can be found in a simpler way [24, p. 398].

2.4.4 Mapping a Discrete-Time into a Continuous-Time Frequency

When analyzed in the time-domain, a continuous-time signal or system can be converted into their discrete-time equivalents by sampling the corresponding time functions at a rate greater than the Nyquist rate. This process is relatively easy, since the time-domain samples can be derived from a given discrete-time sequence by considering them spaced by T_s , the sampling interval.

In the frequency domain, we must also be able to map the discrete-time results into useful continuous-time results. For example, besides considering only the frequency range $[0, 2\pi)$ or $[-\pi, \pi)$ in the frequency response $H(\Omega)$ of a discrete-time system, we should also be able to convert the values of the frequency Ω (in radians) into values of the ordinary frequency f , in hertz. To do so, assume that the continuous-time signals $x(t)$ and $y(t)$ are sampled at a rate $1/T_s$. Assume also that $x(t)$ is a complex exponential. Then we have

$$x(t) = e^{-j2\pi f_0 t} \Rightarrow x[n] = x(nT_s) = e^{-j2\pi f_0 nT_s}. \quad (2.122)$$

Using the eigenfunction property, from (2.116) we can write

$$y(t) = H_c(f_0)e^{j2\pi f_0 t} \Rightarrow y(nT_s) = H_c(f_0)e^{j2\pi f_0 nT_s}. \quad (2.123)$$

The eigenfunction property applied to the discrete-time system is given by

$$y[n] = H_d(\Omega_0)e^{j\Omega_0 n}. \quad (2.124)$$

In (2.123) and (2.124) we have used the subscripts “c” and “d” in $H_c(f_0)$ and $H_d(\Omega_0)$ to identify the continuous-time and discrete-time frequency responses at $f = f_0$ and $\Omega = \Omega_0$, respectively.

By sampling the signal $y(t)$ and taking into account that we want the equality between $y[n]$ and $y(nT_s)$, from (2.123) and (2.124) we obtain

$$y(nT_s) = y[n] \Rightarrow H_c(f_0)e^{j2\pi f_0 nT_s} = H_d(\Omega_0)e^{j\Omega_0 n}. \quad (2.125)$$

Since $H_c(f_0)$ must be equal to $H_d(\Omega_0)$ in (2.125), we must have

$$2\pi f_0 T_s = \Omega_0. \quad (2.126)$$

Then, for any value of f and Ω , the following mapping rule applies:

$$f = \frac{\Omega}{2\pi T_s}. \quad (2.127)$$

Example 2.2 – Consider that a continuous-time sample-function $x(t)$ of a white Gaussian noise is sampled at a rate $1/T_s = 100$ samples per second, resulting in the

discrete-time independent and identically-distributed (i.i.d.) sequence $\{x[n]\}$. These signals are shown in Fig. 2.17 for a 1 second interval.

Now consider a linear time-invariant discrete-time system whose impulse response is given by

$$h[n] = \frac{1}{M} \sum_{k=0}^{M-1} \delta[n - k]. \quad (2.128)$$

This impulse response is plotted in Fig. 2.18 for $M = 10$. A system with impulse response given by (2.128) is one of the variants of the *moving average filter*, a device that smooths the output sequence as compared to the input by forming each output sample from the average of the present and the $M - 1$ preceding input samples, that is,

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n - k] = \frac{1}{M} \sum_{k=0}^{M-1} x[n - k]. \quad (2.129)$$

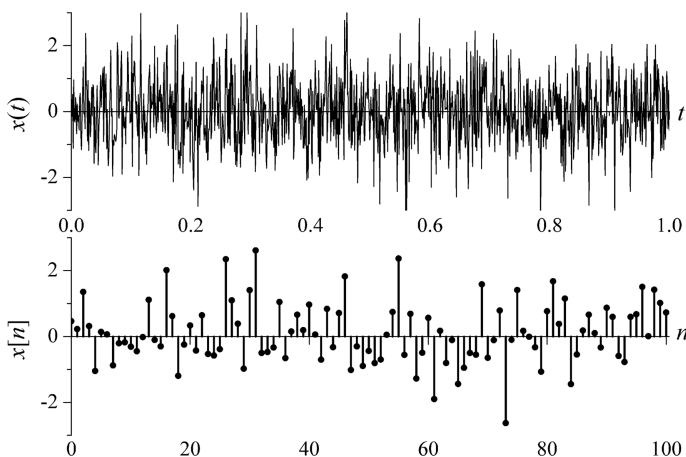


Fig. 2.17 Discrete-time signal $x[n]$ generated by sampling the continuous-time signal $x(t)$

The signal at the output of the moving average filter, as it is determined via (2.129), is plotted in Fig. 2.19. Note that, in fact, this filter smooths the signal. It is a low-pass filter, as we shall see from its frequency response. Note also that the filter has produced a time-correlated sequence, that is, each value in the sequence is not anymore independent from the previous ones.

From (2.98) and (2.128), the frequency response of the moving average filter can be determined from

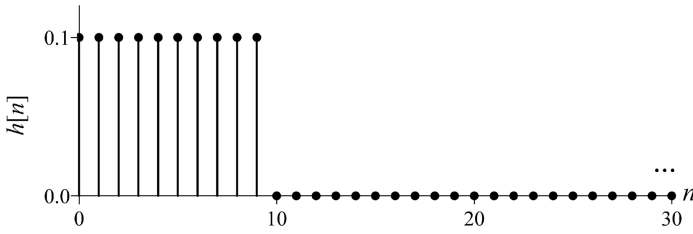


Fig. 2.18 Impulse response $h[n]$ of a moving average filter for $M = 10$

$$H(\Omega) = \sum_{n=-\infty}^{\infty} h[n]e^{-j\Omega n} = \frac{1}{10} \sum_{n=0}^9 e^{-j\Omega n} \sum_{k=0}^9 \delta[n-k] = \frac{1}{10} \sum_{n=0}^9 e^{-j\Omega n}. \quad (2.130)$$

This frequency response is plotted in Fig. 2.20 in terms of its magnitude and phase. Note that, as mentioned before, the frequency response of a discrete-time system is periodic with period 2π . Then, in Fig. 2.20 we must consider only the portion of the graphs between $-\pi$ and $+\pi$. Note also that, in fact, the moving average filter is a low-pass filter.

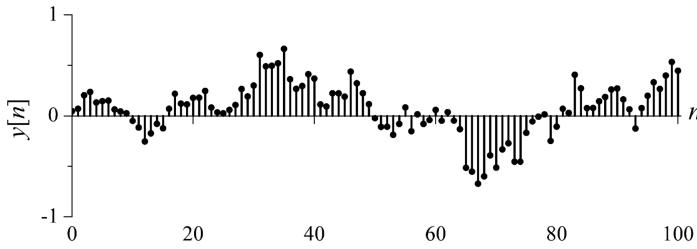


Fig. 2.19 Filtered discrete-time signal $y[n]$

Now, let us map the frequency response of the moving average filter according to the sampling rate given at the beginning of this example. From (2.127) we have

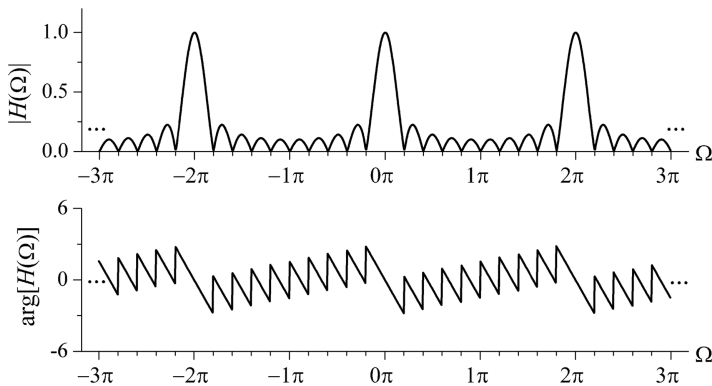


Fig. 2.20 Frequency response of the moving average filter

$$f = \frac{\Omega}{2\pi T_s} \Rightarrow f = \frac{\Omega}{2\pi \times 0.01} = \frac{50\Omega}{\pi}. \quad (2.131)$$

Then, the first spectral null of the magnitude response given in Fig. 2.20 corresponds to the frequency

$$f_{\text{null}} = \frac{50\Omega_{\text{null}}}{\pi} = \frac{50 \times \pi/5}{\pi} = 10 \text{ Hz}. \quad (2.132)$$

Note that this result is consistent with an equivalent continuous-time system, since the duration of the impulse response $h[n]$ is $T = 10 \times T_s = 0.1$ s and the first null is happening at $f = 1/T$.

Simulation 2.5 – Moving Average Filter



File – CD drive: \Simulations\Signals\Moving average.vsm.

Default simulation settings: Frequency = 1,000 Hz; End = 1 second.
Signal source: fixed noise sample-function. Moving average filters:
10-tap.

This experiment aims at complementing the previous example. We have three selectable signal sources: a fixed noise sample-function identical to that shown in the upper part of Fig. 2.17, a random noise sample function that changes at each simulation run and a variable-mean random noise sample-function. The selected signal source is sampled at 100 samples per second and the resulting sequence is applied to a moving average filter. This filter has a configurable impulse response length, which is associated to its number of taps, according to the structure shown in Fig. 2.21. In this figure z^{-1} is a unit-sample delay and $h_n, n = 0, 1, \dots, M - 1$ is a shorthand notation for $h[n]$. Three filter lengths are implemented in this simulation: 10-tap, 20-tap and 40-tap.

At the lower part of the simulation worksheet a unit impulse is applied to an identical moving average filter, aiming at analyzing its frequency response.

Open and run the simulation file indicated in the header using its default settings and observe the time plot. Note that the signals are equal to those shown in Fig. 2.17 and Fig. 2.19.

Now observe the frequency plot and compare the estimated frequency response magnitude with Fig. 2.20. Note that they are plotted in different scales, but are related to the same frequency response. Observe also that it is periodic with period 100 Hz, a value that, from (2.131), corresponds to 2π rd. Note also that the first spectral null is happening at $f = 10$ Hz, as computed in (2.132).

In the signal source block, select the random noise sample-function and, while observing the time plot, run the simulation. This situation was created only to permit that you visualize a more dynamic operation of the moving average filter. Still

using this signal source, change the number of taps of the moving average filter and observe the filtered sequence. Note that the smoothing effect of the filter becomes more evident as the length of its impulse response increases (which means that its cutoff frequency decreases). Note also that the filtered sequence is progressively approaching an all-zero sequence, which is consistent with the fact that the input signal has a zero mean.

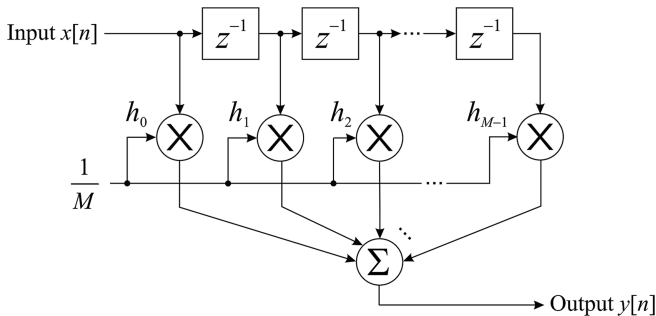


Fig. 2.21 Block diagram of a moving average filter

Now select the variable-mean noise sample-function and increase the simulation end time to 5 seconds. Increase also the number of taps of the filter to 40. Run the simulation while observing the time plots. Note that the input signal clearly represents a variable mean signal. Note also from the filtered sequence that the filter has smoothed-out the noisy component in the input signal, retaining its mean. This is one of the applications of the moving average filter: to extract the variation of the mean value of a random signal.

Finally, analyze the frequency response of the 20-tap and the 40-tap filters. Note that the response becomes narrower as the filter length increases. In other words, an increased filter length increases the duration of the impulse response and, as a consequence, reduces the cutoff frequency of the filter. Try to make calculations concerning the first spectral null and compare your results with the simulation results.

Explore inside the individual blocks. Try to understand how they were implemented. Create and investigate for yourself new situations and configurations of the simulation parameters and try to reach your own conclusions.

2.4.4.1 A Note on Discrete-Time Filtering

The diagram shown in Fig. 2.21 is a generic representation of a *finite impulse response* (FIR) discrete-time system, known as a *transversal filter*. By using different lengths and the *tap coefficients* $\{h_n\}$ we are able to implement a variety of filters with a variety of frequency responses. The moving average filter considered before is just an example of a FIR filter.

A system whose impulse response is nonzero for infinite time is said to be an *infinite impulse response (IIR)* system. The example below considers an IIR filter.

Example 2.3 – A low-pass RC is one of the simplest analog filters. Its frequency response $H(f)$ and the -3 dB cutoff frequency f_0 are given respectively by:

$$H(f) = \frac{1}{1 + j2\pi fRC} \quad \text{and} \quad (2.133)$$

$$f_0 = \frac{1}{RC},$$

where R is the value of the resistor, in ohms (Ω), and C is the value of the capacitor, in farads (F). Figure 2.22 shows the RC filter and the squared magnitude of its frequency response.

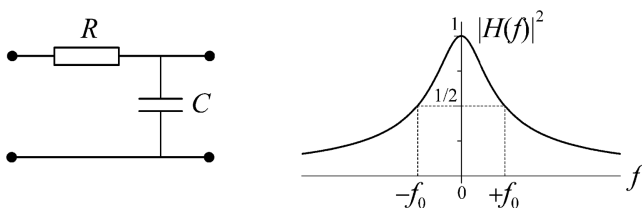


Fig. 2.22 A low-pass RC filter and the squared magnitude of its frequency response

One of the techniques for converting an analog filter into its discrete-time equivalent is called *bilinear transformation* [32, p. 287]. Using this technique, let us construct a discrete-time low-pass filter equivalent to the analog RC filter, for a -3 dB cutoff frequency of $\pi/4$ rad. This frequency corresponds to $f = 0.828/T_s$ Hz, a value that does not follow the mapping given by (2.127) due to the *frequency warping* [32, p. 292] caused by the bilinear transformation method. We omit further details of this conversion method and present the final frequency response of the *RC-equivalent* discrete-time filter:

$$H(\Omega) = \frac{0.293(1 + e^{-j\Omega})}{1 - 0.414e^{-j\Omega}}. \quad (2.134)$$

Note that $H(0) = 1$ and $|H(\pi/4)|^2 = 1/2$, as desired. The diagram of this IIR filter is presented in Fig. 2.23. Note that it is quite different from a transversal FIR filter, since it is the result of the bilinear transformation applied to an analog RC filter. Different transformation methods can lead to different filter structures.

It is left as an exercise to the reader to find that the difference equation that describes the input-output relationship of this filter is given by

$$y[n] = 0.293x[n] + 0.293x[n-1] + 0.414y[n-1]. \quad (2.135)$$

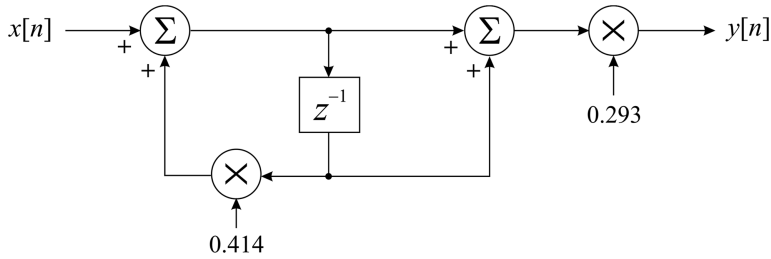


Fig. 2.23 An *RC-equivalent* IIR discrete-time low-pass filter

A *hint* on the above exercise: find how the frequency response given in (2.134) can be converted into a transfer function in the Z-domain. Then, find a general expression for the difference equation of an IIR filter. Finally, find how to write the input-output difference equation based on the Z-domain transfer function.

The project of a digital filter has several approaches, depending on the available initial data and the final application. Some of these approaches are:

- Approximation of a discrete-time filter from an analog filter.
- Discrete-time filter design from a given frequency response mask.
- Choice of available FIR or IIR filters and configuration of their parameters.

The first and second approaches demand a deep knowledge on FIR and IIR filter structures and design criteria. They also demand some specific mathematical background, mainly based on sampling theory and manipulation of discrete-time signals and systems, Fourier, Laplace and Z-transforms and solution of differential and difference equations.

The third approach is the simpler one, since what we need is to choose a given discrete-time filter available in the literature to match our specifications. VisSim/Comm users normally adopt this approach, since several discrete-time filter blocks are already available and the user only needs to find the best one for its application. Nevertheless, the deeper the knowledge of the user in what concerns digital filters, the greater will be his ability to find and configure a better filter. VisSim/Comm also permits that user-defined filters are implemented from “.DAT” files. Both impulse and magnitude responses can be stored in a file and synthesized by the so-called “file FIR filters”. In this case, the first two approaches listed above can be used for a proper design of these filters.

There are also several computer tools for developing digital filters. National Instruments has an Add-on to its *LabVIEW*⁶ software, called *Digital Filter Design*

⁶ LabVIEW is a trademark of National Instruments Corporation. It is a graphical developing software with capabilities of a programming language, data acquisition and analysis, mainly used for development of scalable tests, measurement, and control applications.

Toolkit [22]. It is a powerful tool that extends *LabVIEW* with functions and interactive tools for design, analysis, and implementation of digital filters.

2.5 Complex Representation of Signals and Systems

The complex representation of signals and systems is of great importance for the study of communication systems. In summary, its attributes are:

- It is an alternative representation of signals and systems that, in most cases, reduces the complexity of mathematical treatments.
- It provides an alternative way of representing a passband signal without having to consider the carrier frequency. This is particularly useful in simulations, since the sampling rate can be reduced.
- The power spectral density of a passband signal can be obtained from its complex representation. In most cases, this will lead to a significantly simpler mathematical treatment.

We start by reviewing the concept of the Hilbert transform, from which the complex representation is built. Following [10] and [29], let $x(t)$ be a signal with Fourier transform $X(f)$. The Hilbert transform of $x(t)$ and the corresponding inverse transform are defined respectively by

$$\begin{aligned}\hat{x}(t) &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau = \int_{-\infty}^{\infty} x(\tau) \frac{1}{\pi(t - \tau)} d\tau \quad \text{and} \\ x(t) &= -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{x}(\tau)}{t - \tau} d\tau.\end{aligned}\tag{2.136}$$

In (2.136) we can identify that the Hilbert transform of $x(t)$ is the convolution between $x(t)$ and the function $1/\pi t$, whose Fourier transform is

$$\mathfrak{F}\left(\frac{1}{\pi t}\right) = -j \operatorname{sgn}(f),\tag{2.137}$$

where $\operatorname{sgn}(f)$ is the *sign* function or *signum* function already defined in (2.85).

By recalling that a convolution in the time domain corresponds to a multiplication in the frequency domain, then we can write

$$\hat{X}(f) = -j \operatorname{sgn}(f) X(f).\tag{2.138}$$

From (2.138) we see that the Hilbert transform of $x(t)$ corresponds to a phase shift of -90 degrees for the positive frequencies of $X(f)$ and to a phase shift of $+90$ degrees for the negative frequencies.

Another important definition is the *analytic signal* or *pre-envelope* of the signal $x(t)$, which is given by

$$x_+(t) = x(t) + j\hat{x}(t), \quad (2.139)$$

and from where, using (2.138) and the definition of the *signum* function given in (2.85), we obtain

$$X_+(f) = X(f) + \text{sgn}(f)X(f) = \begin{cases} 2X(f), & f > 0 \\ X(0), & f = 0 \\ 0, & f < 0. \end{cases} \quad (2.140)$$

Now, consider a passband signal $x(t)$ whose bandwidth is essentially confined in $2B$ Hz and is small compared to the carrier frequency f_c . According to (2.140), the analytic spectrum $X_+(f)$ is centered about f_c and contains only positive frequency components. Then, using the frequency-shifting property of the Fourier transform we can write

$$x_+(t) = \tilde{x}(t) \exp(j2\pi f_c t), \quad (2.141)$$

where $\tilde{x}(t)$ is a low-pass signal called the *complex envelope* of the signal $x(t)$.

Since $x_+(t)$ is a passband signal, we can determine the low-pass signal $\tilde{x}(t)$ by a frequency translation of $x_+(t)$ back to about $f = 0$. Using again the frequency-shifting property of the Fourier transform we can write

$$\begin{aligned} \tilde{x}(t) &= x_+(t) \exp(-j2\pi f_c t) \\ &= [x(t) + j\hat{x}(t)] \exp(-j2\pi f_c t), \end{aligned} \quad (2.142)$$

or, equivalently,

$$x(t) + j\hat{x}(t) = \tilde{x}(t) \exp(j2\pi f_c t). \quad (2.143)$$

Since the passband signal $x(t)$ is the real part of the left side of (2.143), we can obtain the so-called *complex representation* of $x(t)$ as

$$x(t) = \text{Re}[\tilde{x}(t) \exp(j2\pi f_c t)]. \quad (2.144)$$

The complex envelope $\tilde{x}(t)$ can be expressed in the Cartesian form by

$$\tilde{x}(t) = x_I(t) + jx_Q(t), \quad (2.145)$$

where the subscripts I and Q stand for *in-phase* and *quadrature*, respectively. Then, by substituting (2.145) in (2.144) we shall have, after some simplifications,

$$x(t) = x_I(t) \cos(2\pi f_c t) - x_Q(t) \sin(2\pi f_c t). \quad (2.146)$$

Both $x_I(t)$ and $x_Q(t)$ are low-pass signals and are called, respectively, the *in-phase component* and the *quadrature component* of the passband signal $x(t)$. This is why we call $\tilde{x}(t)$ the *equivalent low-pass* version of the passband signal $x(t)$.

Rewriting expression (2.145) in its polar form we obtain another useful representation for the complex envelope:

$$\tilde{x}(t) = a(t) \exp[j\theta(t)], \quad (2.147)$$

from where, using (2.144), we obtain

$$\begin{aligned} x(t) &= \operatorname{Re}[\tilde{x}(t) \exp(j2\pi f_c t)] \\ &= \operatorname{Re}\{a(t) \exp[j\theta(t)] \exp(j2\pi f_c t)\} \\ &= a(t) \cos[2\pi f_c t + \theta(t)]. \end{aligned} \quad (2.148)$$

In (2.148), $a(t) = |\tilde{x}(t)|$ is the *envelope* of the passband signal $x(t)$, which can be interpreted as the amplitude modulated component of $x(t)$. Also in (2.148), $\theta(t)$ is the *phase* of the passband signal $x(t)$, or the phase-modulated component of $x(t)$.

Simulation 2.6 – Real Versus Complex Simulation



File #1 – CD drive:\Simulations\Signals\Real_simulation.vsm

Default simulation settings: Frequency = 7,200 Hz; End = 833 seconds. Channel SNR: 5 dB.

File #2 – CD drive:\Simulations\Signals\Cplx_simulation.vsm

Default simulation settings: Frequency = 1 Hz; End = 1,000,000 seconds. Channel SNR: 5 dB.

This experiment aims at illustrating the use of the complex representation as a means to speed-up simulations of communication systems. It comprises two simulation files: “Real_simulation.vsm” (file #1) simulates a digital modulator and demodulator (modem) implemented in a way very close to a real system. File “Cplx_simulation.vsm” (file #2) simulates a digital modem using the complex representation of the modulated signal and, thus, it does not use any carrier signal.

Since we still may not be able to completely understand the structure of a digital modem, we focus our attention to the basic aspects sufficient to show the application of the complex representation in a communication system simulation.

Both modems are using one of the simplest digital modulations: the BPSK (binary phase-shift keying). In the case of file #1, the modulated signal $x(t)$ is generated according to

$$x(t) = b(t) \cos(2\pi f_c t), \quad (2.149)$$

where $b(t)$ is a binary random wave associated to the information bits, such that a binary “1” is represented by a rectangular pulse of amplitude $+1$ and duration T_b and a binary “0” is represented by a rectangular pulse of amplitude -1 and duration T_b . The carrier signal responsible for translating the baseband sequence to passband is the cosine function with frequency f_c Hz. Then, a bit “1” generates a transmitted pulse with frequency f_c and zero phase and a bit “0” generates a transmitted pulse with frequency f_c and phase π radians.

Comparing (2.149) with (2.146) we see that $x_I(t) = b(t)$ and $x_Q(t) = 0$. Then, the complex envelope of the BPSK signal is

$$\tilde{x}(t) = a(t). \quad (2.150)$$

The basic difference between the simulation files considered in this experiment is that file #1 generates the real modulated signal according to (2.149) and file #2 generates the “modulated” signal according to (2.150). Since in file #1 we are simulating a 1,200 bit/s transmission with a 1,800 Hz carrier frequency, the simulation frequency must be at least around 6,000 Hz. We are using 7,200 Hz to have four samples per carrier period.

By observing inside the “BPSK modulator” block in file #2, we see that the modulated signal is being generated by a built-in block of VisSim/Comm. In fact, this block generates the transmitted signal using the complex representation given by the term between brackets in (2.144), which is the complex envelope $\tilde{x}(t)$ multiplied by the complex exponential $\exp(j2\pi f_c t)$. From this observation we notice that if the carrier frequency is set to zero, the transmitted signal is the low-pass signal $\tilde{x}(t)$, which is indeed the case in file #2. In other words, the complex representation approach does not need to use any carrier signal. Moreover, the simulation using $\tilde{x}(t)$ does not demand the sequence $b(t)$ to be generated at 1,200 bit/s. As long as we recalibrate the signal-to-noise ratio (SNR) per bit, we can do the simulation at 1 bit/s and, thus, save additional computational resources.

In both approaches the communication channel is an additive white Gaussian noise (AWGN) channel with adjustable SNR per bit.

The understanding of the receiver structures is beyond our reach by this moment, but it suffices to know that both are optimum in the sense that they minimize the probability of a decision error upon the received bits.

Now run file #1 using its default settings. Wait for the end of the simulation and note that 1,000,000 bits were transmitted and around 6,000 were received in error, yielding a bit error rate (BER) of 6×10^{-3} . Run the simulation again and try to compute the approximate simulation time. Now run the simulation file #2 and also try to compute the simulation time. Maybe you will not be able to do this, since the simulation with file #2 is very fast as compared to the simulation using file #1. Note that the second simulation is also generating 1,000,000 bits and producing around 6,000 bit errors, also resulting in a BER of 6×10^{-3} .

The systems in files #1 and #2 are completely equivalent from the point of view of the performance evaluation, but the complex representation approach saves

computational resources and, as a consequence, can lead to faster simulations than the real representation approach.

Vary the SNR in both cases and compare the performances of both systems. Note that they produce essentially the same BER. Create and investigate for yourself some new situations and configurations of the simulation parameters and try to reach your own conclusions.

2.6 The Power Spectral Density Revisited

In this section we revisit the concept of the power spectral density, complementing the presentation initiated in Chap. 1. First we apply the concept to the complex representation of a passband signal. In addition, we study the *periodogram*, one of the simplest and most used techniques for estimating the PSD of a signal. We also address the general expression for the PSD of multilevel signals with arbitrary correlation between pulses, arbitrary mean and arbitrary pulse shape. This last topic will be particularly useful when analyzing the PSD of digital communication signals, mainly in Chaps. 4 and 6.

2.6.1 The PSD of Passband Signals

Let us assume that the Fourier transform of the signal $x(t)$ in (2.144) exists and is exact. If $x(t)$ is a voltage signal, then the magnitude of its Fourier transform will result in a *voltage spectral density*. Then, from (2.144) we obtain

$$X(f) = \Im\{x(t)\} = \int_{-\infty}^{\infty} \{\operatorname{Re}[\tilde{x}(t)e^{j2\pi f_c t}]\} e^{-j2\pi f t} dt. \quad (2.151)$$

Using the identity $\operatorname{Re}[C] = \frac{1}{2}[C + C^*]$ in (2.151) and applying the Fourier transform properties $z^*(t) \Leftrightarrow Z^*(-f)$ and $z(t) \exp(j2\pi f_c t) \Leftrightarrow Z(f - f_c)$, we obtain

$$\begin{aligned} X(f) &= \frac{1}{2} \int_{-\infty}^{\infty} [\tilde{x}(t)e^{j2\pi f_c t} + \tilde{x}^*(t)e^{-j2\pi f_c t}] e^{-j2\pi f t} dt \\ &= \frac{1}{2} [\tilde{X}(f - f_c) + \tilde{X}^*(-f - f_c)]. \end{aligned} \quad (2.152)$$

Then, the power spectral density $S_X(f)$ can be determined as the squared-modulus of $X(f)$, that is,

$$S_X(f) = |X(f)|^2 = \frac{1}{4} [|\tilde{X}(f - f_c) + \tilde{X}^*(-f - f_c)|^2]. \quad (2.153)$$

Using the simplified notation $\tilde{X}(f - f_c) \equiv Z(f)$, and the fact that $|C|^2 = CC^*$, where C is a complex quantity, we can rewrite (2.153) as follows:

$$\begin{aligned}
 S_X(f) &= \frac{1}{4} [|Z(f) + Z^*(-f)|^2] \\
 &= \frac{1}{4} \{ [Z(f) + Z^*(-f)][Z^*(f) + Z(-f)] \} \\
 &= \frac{1}{4} \left[\begin{aligned} &Z(f)Z^*(f) + Z^*(-f)Z(-f) \\ &+ Z(f)Z(-f) + Z^*(-f)Z^*(f) \end{aligned} \right] \\
 &= \frac{1}{4} \left[\begin{aligned} &|Z(f)|^2 + |Z(-f)|^2 \\ &+ Z(f)Z(-f) + Z^*(-f)Z^*(f) \end{aligned} \right].
 \end{aligned} \tag{2.154}$$

By recognizing that $Z(f)$ and $Z(-f)$ are band-limited, passband signals, the products $Z(f)Z(-f)$ and $Z^*(f)Z^*(-f)$ in (2.154) vanish to zero. Going back to the normal notation, we finally obtain

$$\begin{aligned}
 S_X(f) &= \frac{1}{4} [| \tilde{X}(f - f_c) |^2 + | \tilde{X}(-f - f_c) |^2] \\
 &= \frac{1}{4} [S_B(f - f_c) + S_B(-f - f_c)].
 \end{aligned} \tag{2.155}$$

Equation (2.155) represents an important result and states that we can easily obtain the power spectral density $S_X(f)$ of a passband signal by translating the power spectral density $S_B(f)$ of the low-pass equivalent, and its mirror image, to the frequencies f_c and $-f_c$, respectively, and multiplying the result by 1/4.

Example 2.4 – A random binary wave with equally-likely rectangular pulses of duration $T = 0.2$ second and amplitudes ± 1 volt is the baseband signal that represents the information bits in a communication system according to the mapping rule: bit “0” $\rightarrow -1$, bit “1” $\rightarrow +1$. In a process called *modulation*, this binary signal multiplies a sinusoidal carrier with frequency $f_c = 25$ Hz and random initial phase, generating the *modulated* random process

$$Y(t) = X(t) \cos(2\pi f_c t + \Theta), \tag{2.156}$$

where Θ is the random variable representing the random initial phase of the carrier, assumed uniformly-distributed in $(0, 2\pi]$.

Our task here is to determine $S_Y(f)$, the power spectral density of $Y(t)$.

Comparing (2.156) with (2.146), we can conclude that the quadrature component of the resultant passband process $Y(t)$ does not exist. Furthermore, by intuition we would be tempted to conclude that the random phase will not affect the PSD of the resultant process and then, to simplify matters, we would let $\Theta = 0$ and, in light of (2.145), we would say that the equivalent low-pass process is

$$\tilde{S}(t) = X(t). \quad (2.157)$$

Then, the problem of finding the theoretical PSD of $Y(t)$ would be reduced to finding the PSD of $X(t)$ and applying (2.155).

Intuition sometimes works, but a better solution is to do a more formal analysis. We start this analysis by identifying that the phase of the carrier is independent of the random binary wave. In other words, bits generated by the information source are independent of the phase of the carrier that they will modulate. Then, the autocorrelation function of the process $Y(t)$ is determined according to

$$\begin{aligned} R_Y(\tau) &= E[Y(t)Y(t + \tau)] \\ &= E[X(t) \cos(2\pi f_c t + \Theta) X(t + \tau) \cos(2\pi f_c t + 2\pi f_c \tau + \Theta)] \\ &= E[X(t)X(t + \tau)] E[\cos(2\pi f_c t + \Theta) \cos(2\pi f_c t + 2\pi f_c \tau + \Theta)]. \end{aligned} \quad (2.158)$$

Using the identity $\cos(a) \cos(b) = \frac{1}{2} \cos(a - b) + \frac{1}{2} \cos(a + b)$ we obtain

$$\begin{aligned} R_Y(\tau) &= \frac{1}{2} R_X(\tau) E[\cos(2\pi f_c \tau) + \cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)] \\ &= \frac{1}{2} R_X(\tau) \cos(2\pi f_c \tau). \end{aligned} \quad (2.159)$$

Taking the Fourier transform of both sides of (2.159) and recalling that a multiplication in the time-domain corresponds to a convolution in the frequency-domain, we obtain

$$\begin{aligned} S_Y(f) &= \frac{1}{2} \left\{ S_X(f) * \left[\frac{1}{2} \delta(f - f_c) + \frac{1}{2} \delta(f + f_c) \right] \right\} \\ &= \frac{1}{4} [S_X(f - f_c) + S_X(f + f_c)]. \end{aligned} \quad (2.160)$$

Note that our initial intuition was correct (in this case). Our problem is in fact reduced to finding the PSD of $X(t)$ and applying (2.155).

When comparing (2.160) with (2.155) we must remember that the PSD of a real-valued process is an even function of f , so that $S_X(f + f_c) = S_X(-f - f_c)$.

The autocorrelation function and the power spectral density of a random binary wave were already found in Chap. 1, Examples 1.8 and 1.9. For pulses with amplitudes ± 1 , the PSD of $X(t)$ will be given by

$$S_X(f) = \mathfrak{F}\{R_X(\tau)\} = T \frac{\sin^2(\pi f T)}{(\pi f T)^2} = T \text{sinc}^2(f T). \quad (2.161)$$

Finally, with (2.161) in (2.160), the PSD of $Y(t)$ is given by

$$S_Y(f) = \frac{1}{4} [T \text{sinc}^2[(f - f_c)T] + T \text{sinc}^2[(f + f_c)T]]. \quad (2.162)$$

This PSD, in dBm/Hz, is plotted in Fig. 2.24 for $T = 0.2$ s and $f_c = 25$ Hz.

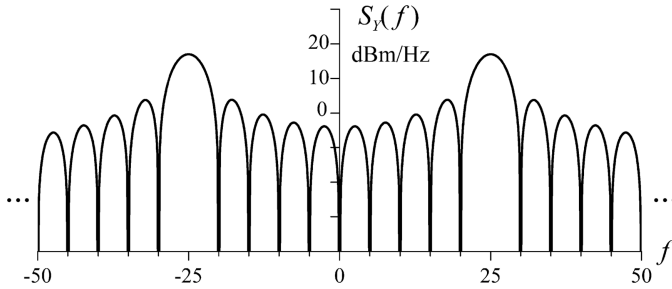


Fig. 2.24 PSD of the random process $Y(t)$ considered in Example 2.4

Anticipating a result that will be considered later on in this chapter, the PSD of a binary antipodal⁷ random wave having equally-likely pulses $\pm g(t)$ can be obtained by dividing the squared-modulus of the Fourier transform of $g(t)$ by the pulse duration, that is,

$$S(f) = \frac{|\mathfrak{F}\{g(t)\}|^2}{T}. \quad (2.163)$$

Note that $|\mathfrak{F}\{g(t)\}|^2$ is the *energy spectral density* of the shaping pulse $g(t)$.

In this example $g(t)$ is a rectangular pulse of duration $T = 0.2$ s and unit amplitude. With the Fourier transform of $g(t)$ obtained from Table 2.4, the PSD of a random sequence of these pulses is given by

$$S_X(f) = \frac{|\mathfrak{F}\{g(t)\}|^2}{T} = \frac{|T \text{sinc}(fT)|^2}{T} = T \text{sinc}^2(fT). \quad (2.164)$$

Note that this result is in agreement with the result in (2.161).

Simulation 2.7 – Estimating the PSD of Passband Signals



File – CD drive: \Simulations\Signals\Passband.PSD.vsm

Default simulation settings: Frequency = 100 Hz; End = 1,000 seconds. Probability of a pulse “+1” in the random binary wave source: 0.5.

⁷ An antipodal signal is composed by a sequence of passband or baseband pulses having shape $g(t)$ and opposite polarities. Then, for example, if an antipodal signal represents a sequence of bits, we may have the mapping rule: bit “0” $\rightarrow -g(t)$; bit “1” $\rightarrow +g(t)$, or vice-versa.

This simple experiment complements Example 2.4, illustrating the estimation of the PSD of a passband signal from the knowledge of the PSD of its equivalent low-pass representation. The computation of the PSD is made by VisSim/Comm using a method called periodogram, a topic that will be covered in more detail in the next subsection. For now it suffices to know that this method computes the PSD as an average of Fourier transform results obtained from finite duration sample functions of the signal under analysis.

A random antipodal binary wave is generated at 5 pulses per second and with configurable probability of a pulse “+1”. This wave modulates a sinusoidal carrier with frequency $f_c = 25$ Hz.

A time plot and a histogram plot are connected to the output of the binary source to permit a better analysis of the random binary signal. The power spectral densities of the binary (baseband) and modulated (passband) signals are estimated through periodograms. The estimation is made via an 8k-point FFT and 12 FFT averages, using a rectangular window.⁸ The results are displayed in dBm/Hz.

Using the default settings, run the simulation and compare the PSD of the binary wave with the one obtained theoretically in Chap. 1, Sect. 1.14.2 (Example 1.9). Compare the PSD of the passband modulated signal with that illustrated in Fig. 2.24. Note that the theoretical and the estimated PSDs are approximately equal to one another. Try to measure the peaks of the binary wave PSD and of the modulated signal. Note that the PSD peak of the modulated signal is approximately 6 dB below the PSD peak of the binary wave, a consequence of the multiplication by $1/4$ in (2.155), since $10 \log(1/4) \cong -6$ dB.

Now run the simulation while observing the time and histogram plots. Note that these plots are in agreement with the configured pulse probability. Make changes in the distribution of the probabilities of the pulses and rerun the simulation, while observing again the time plot and the histogram. Note the correspondence between the observed plots and the configured probability. Now observe the estimated PSDs. Note that discrete spectral components have appeared on the top of the main lobes, a result that is consistent with the change in the mean value of the random wave due to the unbalance in the pulse probabilities. This unbalance corresponds to a DC (zero-frequency) component in the PSD of the baseband random wave and to a discrete component at the carrier frequency in the modulated signal.

The effect of unbalanced bit probabilities in the PSD of communication signals will be covered in more detail in Subsect. 2.6.3. In addition, several examples involving typical communication waveforms will be addressed in Chap. 4, where baseband transmission is discussed in detail.

As an exercise, try to implement modifications in the sinusoidal carrier in order to simulate a random initial phase. After that, investigate the influence of this

⁸ The use of a *rectangular window* means that the signal samples used by the FFT computation are selected from the signal under analysis as if the signal were multiplied by a rectangular function with duration equal to the window length.

initial phase on the estimated PSDs. *Hint*: use the “Sinusoid” block from the “Comm”/“Signal Sources” tabs. It permits the initial phase to be configured.

Explore inside the individual blocks. Try to understand how they were implemented. Create and investigate for yourself new situations and configurations of the simulation parameters and try to reach your own conclusions.

2.6.2 Estimation of the PSD via the Periodogram

If the signal $x(t)$ is a sample-function of the stationary random process $X(t)$, its Fourier transform may not exist. In this case, as we saw in Chap. 1, the spectral content of $X(t)$ is given by its *power spectral density* (PSD), which is obtained from the Fourier transform of the autocorrelation function $R_X(\tau)$ of the random process, that is,

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f\tau} d\tau. \quad (2.165)$$

The Parseval’s theorem [24, p. 221] states that the total energy contained in a waveform integrated across all time is equal to the total energy of the waveform’s Fourier transform integrated across all of its frequency components. This is also known as the Rayleigh’s energy theorem. By defining $X_\Pi(f)$ as the Fourier transform obtained from the sample-function $x_\Pi(t)$, which is $x(t)$ truncated from $-\Pi/2$ to $\Pi/2$, the Parseval’s theorem can be written as

$$\int_{-\Pi/2}^{\Pi/2} x^2(t) dt = \int_{-\infty}^{\infty} |X_\Pi(f)|^2 df, \quad (2.166)$$

where $|X_\Pi(f)|^2$ is called the *energy spectral density* of the energy signal $x_\Pi(t)$. Multiplying both sides of (2.166) by $1/\Pi$ and making $\Pi \rightarrow \infty$ we obtain

$$\lim_{\Pi \rightarrow \infty} \frac{1}{\Pi} \int_{-\Pi/2}^{\Pi/2} x^2(t) dt = \int_{-\infty}^{\infty} \lim_{\Pi \rightarrow \infty} \frac{|X_\Pi(f)|^2}{\Pi} df. \quad (2.167)$$

Note that in the left-side of (2.167) we have the definition of the average power of the sample-function $x(t)$. To obtain the average power of the process $X(t)$ we must take the expected value of both sides of (2.167), yielding [26, p. 174]

$$P_X = \lim_{\Pi \rightarrow \infty} \frac{1}{\Pi} \int_{-\Pi/2}^{\Pi/2} E[X^2(t)] dt = \int_{-\infty}^{\infty} \lim_{\Pi \rightarrow \infty} \frac{E[|X_\Pi(f)|^2]}{\Pi} df. \quad (2.168)$$

Finally, the integrand on the right-side of (2.168) must be the power spectral density of the random process $X(t)$ [10, p. 51], since its integral results in a power. Furthermore, the unit watts/hertz for the PSD is adequate.

In (2.168) we notice that the average power in the process $X(t)$ is given by the time average of its second moment. Nevertheless, if $X(t)$ is stationary $E[X^2(t)]$ will be constant and the total average power in $X(t)$ will be $P_X = E[X^2(t)]$.

From the above discussion we have the following PSD estimation for $X(t)$:

$$S_X(f) = \lim_{\Pi \rightarrow \infty} \frac{1}{\Pi} E[|X_\Pi(f)|^2]. \quad (2.169)$$

If the signal $x(t)$ is deterministic, expression (2.169) can also be used, but without the expectation operation [1, p. 31].

The PSD estimation in (2.169) is called a *periodogram* [21, p. 212; 27, p. 902]. A useful interpretation about it can be stated as follows: the power spectral density of a random process $X(t)$ can be estimated by averaging the squared-modulus of the Fourier transform of windowed sample-functions. For stationary random processes, this averaging can be obtained through time-averaging.

In practice the periodogram estimation of a PSD is accomplished by segmenting a sample-function, sampling each segment, performing a discrete Fourier transform (typically a fast Fourier transform, FFT), applying the squared-modulus, summing the results coming from each segment and dividing the result by the number of segments. This averaging process is sometimes called *smoothing*.

The windowing process is not necessarily rectangular and, in fact, different windowing can reduce *bias* and *leakage* in the estimated PSD [21, p. 216], as we shall see in the next computer simulation.

The periodogram is just one of the many techniques for estimating the power spectral density of a signal. For more information on this technique, see [18, p. 622]. For a detailed analysis on several other techniques for power spectral density estimation, see [21] and [36]. For a condensed material on several aspects of the PSD estimation, see Chap. 12 of [27].

Simulation 2.8 – Periodogram Estimation of the PSD



File – CD drive: \Simulations\Signals\Periodogram.vsm

Default simulation settings: Frequency = 1 Hz; End = 32,768 seconds. Number of FFT averages: 32. Window type: Rectangular. FFT size: 1,024 points (1k). PSD unit: dBm/Hz. 1Ω load.

This experiment is intended to show how the spectral estimation via periodogram is performed. The effects of window size and smoothing are investigated. The experiment also shows the bias effect and one possible solution implemented by changing the window type.

The sample-function to be analyzed is an *autoregressive*, usually denoted by AR(4) [21, pp. 164, 218]. This random process is generated according to

$$x[n] = -\sum_{k=1}^4 a_k x[n-k] + w[n], \quad (2.170)$$

where $x[n]$, $n = 0, 1, \dots, N-1$ is a sample of the autoregressive sequence \mathbf{x} and $w[n]$ is a sample of a normal random variable $\mathcal{N}(0, 1)$, forming the sequence \mathbf{w} . The vector \mathbf{a} is given by

$$\mathbf{a} = [1 \quad -2.7607 \quad 3.8106 \quad -2.6535 \quad 0.9238]^T. \quad (2.171)$$

A vector \mathbf{x} with 32,768 (32k) samples was generated according to (2.170) and the result was stored in a .DAT file. This file is read by VisSim/Comm and the computation of the power spectral density is then performed.

The theoretical power spectral density of the AR(4) sequence, in dBm/Hz and considering a 1Ω load resistance is given by

$$S_X(f) = 10 \log \frac{\left| \sum_{k=0}^4 a_k e^{-j2\pi f k} \right|^{-2}}{0.001} \quad (2.172)$$

This PSD was converted into a .DAT file and also exported to VisSim/Comm, so that it is possible to compare the theoretical and the estimated PSDs, which are shown in Fig. 2.25. The estimated periodograms were obtained from a 1k-point FFT using two different window types. The interpretation of the results will be provided throughout the analysis of this experiment.

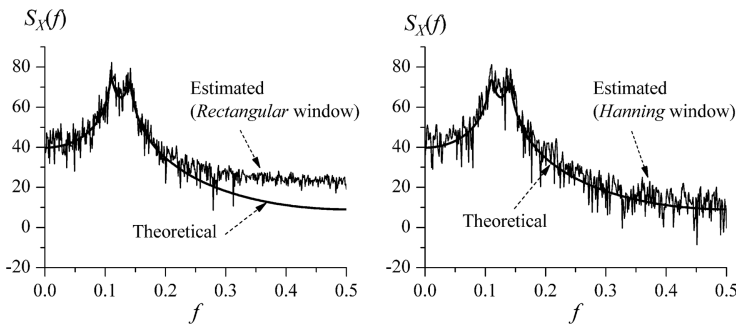


Fig. 2.25 Theoretical and estimated PSDs of the AR(4) signal

The use of the AR(4) signal was motivated due to the fact that it has a high dynamic range and is extensively used for testing PSD estimation techniques [21].

As already mentioned in Chap. 1, the use of the scale in dBm/Hz is motivated by the fact that lower values of the PSD are emphasized to permit measurements. In a linear scale these lower levels perhaps would not be measurable.

Open the simulation file indicated in the header. Note that the sample-function of the AR(4) process is applied to a VisSim/Comm block that estimates the PSD. The result is plotted in dBm/Hz, along with the theoretical PSD given by (2.172). A time plot and an histogram plot are also provided so that you can better analyze the waveform corresponding to the AR(4) signal.

Using the default simulation settings, run the simulation and observe the time plot. Magnify the waveform for a better view. Note that the AR(4) signal seems to be, and in fact is, a filtered white Gaussian process. The histogram plot can confirm this. As a consequence, the theoretical PSD is determined only by the discrete-time filter coefficients $\{a_k\}$, since the PSD of a white noise is constant.

Now observe the estimated PSD and compare it to the theoretical one. Note that there is a strong *bias* for higher frequencies, a typical characteristic of the periodogram that manifests more visibly when the signal has a high dynamic range.

Open the dialog box of the power spectral density analyzer. Note that it is using a relatively short FFT size (1k) and that it is using a smoothing with 32 FFT averages. Note also that the window of a rectangular type is being used. Reduce the number of FFT averages to 1 and rerun the simulation. Note that the variance of the estimated PSD has been increased. This shows to us that the smoothing process is a means for reducing the variance of the periodogram.

Now increase the FFT size to 16k and set the number of FFT averages to 2. Remember that whenever you change the FFT size in the PSD analyzer block, you must set the same value for the theoretical PSD (Spectral Mask block). Run the simulation and note that this increased FFT size has increased the *frequency resolution* of the PSD estimation. Note also that the bias has been slightly reduced. This situation is an example of the increased resolution and reduced bias caused when the FFT size is increased. However, note that the variance of the PSD *is not* affected by the window size [18, p. 622].

If desired, it is possible to save the data corresponding to the PSD plots so that you can use a mathematical tool to compute the variance of the estimated PSDs.

Now, reset the diagram to the default parameters. Run the simulation and observe again the estimated PSD and the bias previously verified. Open the dialog box of the spectral analyzer and change the window type to *Hanning*. Run the simulation and note that the bias has been reduced. This situation shows to us that it is possible to overcome the bias problem, at least by a small amount, using a different windowing. Nevertheless, care must be taken when using windows other than the rectangular, since the gradual decaying of a non-rectangular window is normally accompanied by a *smearing* of the peaks in the estimated PSD.

As an exercise, do some research in order to better understand the various types of window and their main characteristics. Any good book on discrete-time signal processing or digital signal processing (DSP) will probably discuss the subject.

There are many other situations that can be simulated with this experiment. Create and investigate for yourself some new situations and configurations of the

simulation parameters and try to reach your own conclusions. Additionally, try to reproduce the results presented in Fig. 2.25.

2.6.3 The PSD of Baseband Digital Communication Signals

In this subsection we address a general expression for the power spectral density of any baseband digital signals that can be represented in the form

$$x(t) = \sum_{i=-\infty}^{\infty} a_i p(t - iT), \quad (2.173)$$

where $\{a_i\}$ is a sequence with M amplitudes, each of them representing $\log_2 M$ bits, $p(t)$ is a unit-amplitude pulse-shaping function and T is the pulse duration. We call $1/T$ the *signaling rate* or *symbol rate* (or yet *baud rate*), which is measured in symbols per second or *bauds*.

The waveform represented by (2.173) is a form of baseband signaling usually referred to as *M*-ary *Pulse Amplitude Modulation* (*M*-PAM). It is one of the baseband transmission forms that will be covered in detail in Chap. 4, in the more specific context of digital transmission. Additionally, it is associated to the equivalent low-pass representation of several digital modulated signals. Then, determining its PSD is indeed of major importance.

Here we omit some intermediate mathematical deductions and proofs and give special attention to the interpretation and use of the main final results. For a detailed analysis, see for example [6, p. 415; 41, Chap. 2 and Appendix A].

Since the PSD of a signal depends on its autocorrelation function, the correlation between pulses in the sequence $\{a_i\}$ will affect the result. Furthermore, the mean of the sequence $\{a_i\}$ will affect the PSD, since a non-zero mean causes the appearance of discrete components in the frequency spectrum. Additionally, the pulse-shape function $p(t)$ will also play its role in the final spectral shape.

What we shall do in the sequel is to analyze the continuous and discrete parts of the power spectral density of $x(t)$ as a function of the correlation, mean and pulse-shaping properties of the signal. We assume that $x(t)$ is a sample-function of an ergodic random process $X(t)$. This assumption is reasonable for many communication signals, especially if they are not modulated onto a carrier or, if modulated, they are operated through their complex envelope representation [14].

We start by defining the discrete autocorrelation function of the sequence $\{a_i\}$:

$$R_A(k) = E[a_i a_{i+k}]. \quad (2.174)$$

Since the sequence $\{a_i\}$ can have any mean μ_A and variance σ_A^2 , we define the normalized zero-mean and unit-variance sequence

$$\hat{a}_i = \frac{a_i - \mu_A}{\sigma_A}. \quad (2.175)$$

As a consequence, we have a normalized discrete autocorrelation function

$$\hat{R}_A(k) = E[\hat{a}_i \hat{a}_{i+k}] = \frac{R_A(k) - \mu_A^2}{\sigma_A^2}. \quad (2.176)$$

This autocorrelation can be estimated through the discrete-time average

$$\hat{R}_A(k) \cong \frac{1}{n} \sum_{j=0}^{n-1} \hat{a}_j \hat{a}_{(j+k) \bmod n}. \quad (2.177)$$

In (2.177) the modulo- n operation can be avoided if the number of points in the sequence $\{\hat{a}_i\}$ is at least $N = 2n - 1$. Moreover, for a good estimation n must be as large as possible. In fact (2.177) corresponds to the exact computation of the autocorrelation function for any periodic sequence (pseudo-random sequences, for example). Furthermore, for large values of n , the circular-shift caused by the modulo- n operation does not cause large estimation errors for small values of the lag k .

Now let us define a spectral weighting function $\hat{S}_A(f)$, given by the Fourier transform of the normalized *continuous-time* autocorrelation function of the sequence $\{\hat{a}_i\}$, which can be written in terms of the *discrete-time* autocorrelation function, as shown by:

$$\begin{aligned} \hat{S}_A(f) &= \int_{-\infty}^{\infty} \hat{R}_A(\tau) e^{-j2\pi f \tau} d\tau \\ &= \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \hat{R}_A(k) \delta(\tau - kT) e^{-j2\pi f \tau} d\tau \\ &= \sum_{k=-\infty}^{\infty} \hat{R}_A(k) \int_{-\infty}^{\infty} \delta(\tau - kT) e^{-j2\pi f \tau} d\tau \\ &= \sum_{k=-\infty}^{\infty} \hat{R}_A(k) e^{-j2\pi f kT}. \end{aligned} \quad (2.178)$$

This spectral weighting function is associated to the continuous part of the spectrum of $x(t)$, since it depends on the autocorrelation function of the zero-mean and unit-variance sequence $\{\hat{a}_i\}$.

Then, for the general case where it is possible that the sequence $\{a_i\}$ in (2.173) exhibits correlation between pulses and have a non-zero mean, the power spectral density of the signal $x(t)$ defined in (2.173) is given by [6, p. 418; 41, p. 577]

$$S_X(f) = \frac{|P(f)|^2}{T} \sigma_A^2 \hat{S}_A(f) + \left(\frac{\mu_A}{T}\right) \sum_{i=-\infty}^{\infty} \left|P\left(\frac{i}{T}\right)\right|^2 \delta\left(f - \frac{i}{T}\right). \quad (2.179)$$

Note that the second term in the right-hand side of (2.179) is the discrete part of the spectrum. For a non-zero mean of the sequence $\{a_i\}$ in (2.173), this term will correspond to discrete components in harmonics of the signaling rate $1/T$. It will vanish if the sequence $\{a_i\}$ has zero-mean, that is $\mu_A = 0$.

The first term in the right-hand side of (2.179) is influenced by the pulse-shaping function $p(t)$ and by the correlation properties of the sequence $\{a_i\}$. Note that for uncorrelated pulses, the normalized autocorrelation function in (2.176) will be given by a unity impulse at $k = 0$. With this result in (2.178) we shall have a unitary weighting function $\mathfrak{S}_A(f)$ and the power spectral density will be determined only by the pulse-shaping function $p(t)$ and the variance σ_A^2 of the sequence $\{a_i\}$.

Example 2.5 – Let us apply expression (2.179) to the analysis of the PSD of the random binary wave with uncorrelated, rectangular and equally-likely pulses $\{a_i\} \in \pm A = \pm 1$ already considered in Example 2.4.

Since the pulses are equally-likely, $\mu_A = 0$. The variance σ_A^2 is given by $E[(a_i - \mu_A)^2] = E[a_i^2] = A^2 = 1$.

Additionally, since the pulses are uncorrelated, the spectral weighting function $\mathfrak{S}_A(f) = 1$. Then, with these results in (2.179) we obtain

$$S_X(f) = \frac{|P(f)|^2}{T}. \quad (2.180)$$

Since $p(t)$ is a rectangular pulse of unit amplitude and duration T we can apply the already known Fourier transform, yielding

$$S_X(f) = \frac{T^2 \text{sinc}^2(fT)}{T} = T \text{sinc}^2(fT). \quad (2.181)$$

This is indeed the same result obtained in (2.161) and (2.164), demonstrating that the power spectral density of a zero-mean random sequence of uncorrelated pulses can be obtained by dividing the energy spectral density of the pulse-shaping function by the pulse duration T and multiplying the result by the variance (in this case equal to the mean square value) of the sequence $\{a_i\}$, that is,

$$S_X(f) = \frac{|P(f)|^2}{T} \sigma_A^2. \quad (2.182)$$

Comparing (2.182) with (2.163), observe that (2.163) applies only to binary antipodal signals with pulses $\pm g(t)$ and is a particular case of (2.182) if we use $G(f) = P(f)\sigma_A$, that is, $g(t) = p(t)\sigma_A$. Nevertheless, (2.182) is applicable to any

zero-mean random sequence of uncorrelated multilevel pulses, as long as the shape of the pulses, $p(t)$, is the same.

In [28, p. 530] and [37] the spectral characteristics of digital signals are analyzed from the perspective of Markov chains. With this approach, signals are represented by a set of states that happen with specified probabilities and with specified transition probabilities between states. This analysis is particularly useful for determining the power spectral density of signals with memory, that is, signals in which the correlation between one or more successive pulses are intrinsic to the signal generation. Examples of such signals are some line codes that will be considered at the beginning of Chap. 4.

2.7 The Bandwidth of Communication Signals

The information provided by the power spectral density of a random process or a deterministic signal can be complemented if we associate to it some measure of the frequency range in which a given *portion* of the signal power is concentrated on. This measure is usually referred to as the *bandwidth* of the signal.

There are several definitions for bandwidth and they differ basically in the way that the *portion* of the signal power is defined. Usual bandwidth measurements are the *absolute bandwidth*, the -3 dB or *half-power bandwidth*, the *equivalent noise bandwidth*, the *root mean square (rms) bandwidth* and the *null-to-null* or *main lobe bandwidth*. Bandwidth requirements are also commonly defined in terms of a *spectral mask*, which establishes limits for PSD levels. This bandwidth measurement is commonly used by regulatory agencies to impose limits to the radiated power spectrum of communication systems.

In what follows we shall briefly discuss each of these bandwidth definitions.

2.7.1 Absolute Bandwidth

The absolute bandwidth is a measure of the frequency range in which the PSD of a signal is non-zero. The binary random wave considered in Example 2.4 is an example of infinite absolute bandwidth.

Care must be taken when considering the absolute bandwidth of signals defined according to (2.173). As we saw, the PSD of such signal is determined by (2.179), where $P(f)$ is the Fourier transform of the shaping-pulse $p(t)$. As long as $p(t)$ is confined in some interval, no matter how long, $P(f)$ will have infinite bandwidth and, as a consequence, $S_X(f)$ will also have infinite bandwidth. From another point of view, any confined spectrum $P(f)$ is associated to an unlimited-time pulse $p(t)$.

2.7.2 -3 dB (or Half Power) Bandwidth

As the name indicates, the -3 dB bandwidth is the range of frequencies comprised between the frequency values in which the PSD of a signal falls 3 dB related to its peak. If the PSD has multiple occurrences of -3 dB points, we can consider the bandwidth of the -3 dB frequency values that are mostly far apart.

2.7.3 Equivalent Noise Bandwidth

At the end of Chap. 1 we defined the equivalent noise bandwidth B as the bandwidth of an ideal filter that produces the same noise power as the filter representing a given system's response, when a white noise is applied to both. Translating this definition to the context of power spectral density, we can define the equivalent noise bandwidth as the bandwidth of a brick-wall-shaped PSD that has the same area of the PSD of a given process $X(t)$. Making use of the results in Chap. 1, we then have

$$2 \int_0^\infty S_X(f) df = S_X(0)2B \Rightarrow B = \frac{\int_0^\infty S_X(f) df}{S_X(0)}. \quad (2.183)$$

2.7.4 Root Mean Square (rms) Bandwidth

The concept behind the *rms* bandwidth comes from the notion of the standard deviation as a measure of the dispersion of a probability density function.

If the PSD is integrable, we can define a normalized PSD such that the resultant area is made equal to 1. From this PDF-like power spectral density we can compute the square root of the second moment and associate the result to a measure of the dispersion of the SPD. Mathematically, for a baseband real signal we have

$$B = \sqrt{\int_{-\infty}^{\infty} f^2 \left(\frac{S_X(f)}{\int_{-\infty}^{\infty} S_X(u) du} \right) df}. \quad (2.184)$$

For a passband signal we can apply the same reasoning. First we define the mean frequency μ_F as

$$\mu_F = \int_0^\infty f \left(\frac{S_X(f)}{\int_0^\infty S_X(u) du} \right) df. \quad (2.185)$$

Then we ignore the negative part of $S_X(f)$ and translate the positive part to $f = 0$. As a result we have a baseband PSD $S_Y(f)$. Finally we apply (2.184), which is multiplied by 2 to reflect the double deviation about the mean in the passband spectrum. Mathematically we have:

$$\begin{aligned}
S_Y(f - \mu_F) &= S_X(f)u(f), \\
S_Y(f) &= S_X(f + \mu_F)u(f + \mu_F) \quad \text{and} \\
B &= 2 \sqrt{\int_{-\infty}^{\infty} f^2 \left(\frac{S_Y(f)}{\int_{-\infty}^{\infty} S_Y(u)du} \right) df},
\end{aligned} \tag{2.186}$$

where $u(f)$ is the unit-step function in the frequency domain. Note that if the positive part of $S_X(f)$ is symmetric about the carrier frequency f_c , the mean frequency will (as it should) be equal to the carrier frequency.

Instead of working with $S_Y(f)$ we can determine the *rms* bandwidth using

$$B = 2 \sqrt{\int_0^{\infty} (f - \mu_F)^2 \left(\frac{S_X(f)}{\int_0^{\infty} S_X(u)du} \right) df}. \tag{2.187}$$

2.7.5 Null-to-Null or Main Lobe Bandwidth

The null-to-null or main lobe bandwidth is a usual measure of the frequency range occupied by a signal whose power spectral density is of the form, or at least is similar in shape to a “sinc” function.

As an example, refer to the PSD in Fig. 2.24. Note that the null-to-null bandwidth is located from 20 to 30 Hz. This is indeed the portion of the signal that concentrates most of the signal power, typically more than 90%. For example, the proportion of the signal power in the main lobe of $S_Y(f)$ in (2.162) is

$$\int_{20}^{30} S_Y(f)df \bigg/ \int_0^{\infty} S_Y(f)df \cong 0.903. \tag{2.188}$$

For a baseband signal, the null-to-null bandwidth lose a little bit its meaning, since the negative half part of the main lobe does not exist. In this case the term “main lobe” or “first null” bandwidth is more adequate, as it measures the frequency band from 0 to the first spectral null.

The power concentration is also valid for the baseband spectrum, as shown below for the PSD $S_X(f)$ in (2.161):

$$\int_0^{1/T} S_X(f)df \bigg/ \int_0^{\infty} S_X(f)df \cong 0.903. \tag{2.189}$$

It is worth mentioning that the term “null” not necessarily refers to a zero value of the PSD. The term also applies to situations where deep notches exist in the spectral content of a signal.

Summarizing the above discussion, Fig. 2.26 shows the bandwidth definitions presented so far, considering a passband signal as a case study. The representation of the equivalent noise and the *rms* bandwidths with smaller values than the

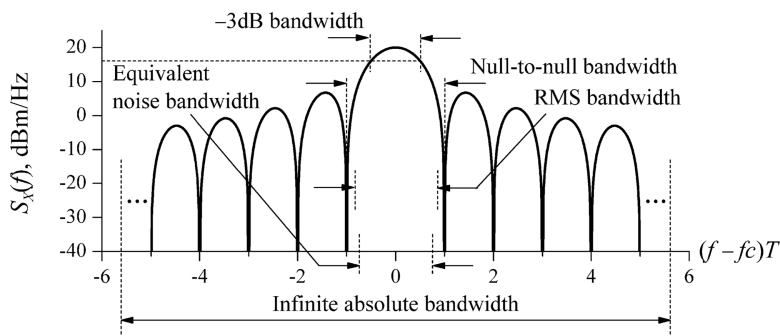


Fig. 2.26 Several bandwidth definitions

null-to-null bandwidth is only illustrative, their exact values depending of the specific signal under analysis.

2.7.6 Spectral Mask

The spectral mask is not a measure of bandwidth, at least in the sense of the other measures described so far. Instead, it establishes upper (and sometimes lower) bounds on the spectral characteristics of a signal. The spectral mask is commonly used by regulatory agencies to impose limits on the radiated power of wireless communication systems.

In Fig. 2.27 the concept of the spectral mask is illustrated. The mask consists of a set of lines that establish the limits for the spectrum of a transmitted signal. Although these limits can be specified in several ways, a normal unit is the “dBc”, which means that they are measured in dB relative to the signal power at the carrier frequency. By using this unit of measurement, transmitters with different average powers can be assessed by using the same spectral mask.

The spectral mask is also used to establish lower and upper bounds simultaneously. In this case the spectral characteristics of the signal or system under analysis

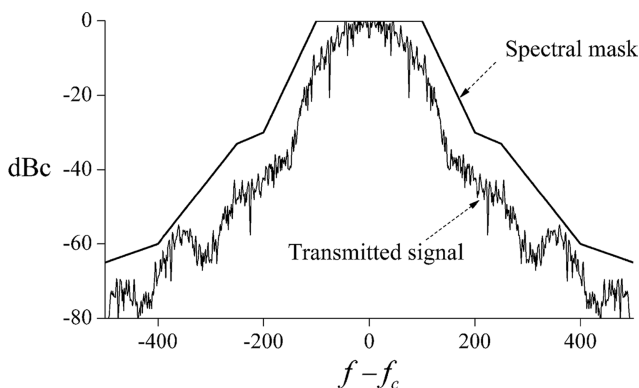


Fig. 2.27 The concept of spectral mask

must fit in-between the imposed limits. This type of spectral mask is particularly useful to set limits in the spectral characteristics of filters.

Just as a complement, a mask can also be defined in the time-domain, having typically lower and upper limits. This type of mask is normally used to set limits in pulse shapes and the like.

2.8 Summary and Further Reading

In this chapter we have presented the main concepts related to communication signals and systems. We have defined some common signals and studied their properties. In the sequel we addressed the time and frequency representation of such signals through the Fourier analysis. The sampling theory was then discussed and linear systems were put into focus, with emphasis in the class of linear, time-invariant (LTI) systems. The complex representation of signals and systems were also considered and the chapter ended with an analysis of the power spectral density and the bandwidth of signals, complementing the corresponding study started in Chap. 1.

The literature is vast in the topics covered in this chapter. However, some classical books and other non-classical but invaluable ones can be highlighted:

- Signals and linear systems, sampling theory and Fourier analysis are treated in detail by H. P. Hsu [13], A. V. Oppenheim, A. S. Willsky and S. H. Nawab [24], S. Haykin and B. V. Veen [11] and B. Girod, R. Rabenstein and A. Stenger [7].
- The discrete-time signal processing area is well covered in the classic book by A. V. Oppenheim and R. W. Schaffer [23] and in the book by V. K. Madisetti and D. B. Williams (editors) in [20].
- In what concerns the power spectral estimation, a deep mathematical treatment is given by R. M. Howard in [12], and an applied material is presented in Chap. 5 of the book by D. G. Manolakis, V. K. Ingle and S. M. Kogon [21]. For a condensed material on several aspects of PSD estimation, see Chap. 12 of the book by J. G. Proakis and D. G. Manolakis [27]. A more complete covering of the subject is given by P. Stoica and R. L. Moses in [36]. In the book by J. G. Proakis and M. Salehi [28, p. 530] and in the report by R. C. Titsworth and L. R. Welch [37], the spectral characteristics of digital signals are analyzed from the perspective of Markov chains. This analysis is particularly useful for determining the power spectral density of signals with memory.
- For a broad discussion on digital filtering, see the works by V. K. Madisetti and D. B. Williams (editors) [20], C. B. Rorabaugh [32] and B. A. Shenoi [33].

2.9 Additional Problems

Several simulation-oriented problems were already proposed in the body of the simulation work plans throughout the chapter. Most of the following additional

propositions are also based on simulation or are research-oriented, requiring some sort of additional reading. A number of exercises involving conventional calculations can be found in the references cited throughout the chapter.

1. Determine the Fourier transform of the following signals:

$$x_1(t) = \begin{cases} K \exp \left[-\frac{(t-T/2)^2}{T^2/8\pi} \right], & 0 \leq t \leq T \\ 0, & \text{otherwise} \end{cases}$$

$$x_2(t) = \begin{cases} \sin \left(\frac{\pi t}{T} \right), & 0 \leq t \leq T \\ 0, & \text{otherwise} \end{cases}$$

$$x_3(t) = \left[\frac{\sin(\pi t/T)}{\pi t/T} \right] \frac{\cos(\pi \alpha t/T)}{1 - (2\alpha t/T)^2}$$

$$x_4(t) = \sqrt{\frac{2B^2\pi}{\ln 2}} \exp \left(-\frac{2B^2\pi^2 t^2}{\ln 2} \right).$$

2. Determine the power spectral density of a 4-PAM signal with independent pulses with rectangular shape, duration T and amplitudes $\pm A$ and $\pm 3A$ volts.
3. Find the specifications of the compact disc digital audio technology in what concerns sampling and quantization and determine the signal-to-quantization noise ratio that can be achieved. Repeat the process considering the Blu-ray disk technology.
4. Do some research and find applications of the moving-average filter in the context of wireless channel characterization. *Hint*: seek for topics related to the analysis of short-term and long-term propagation characteristics of the channel during the so-called channel-sounding processes.
5. Revisit Example 2.3 and prepare a dissertation aiming at detailing all steps involved in the design of a digital filter from its analog counterpart.
6. Do a research on a method for power spectral density estimation different from the periodogram and describe it in details. If possible, implement the method using VisSim/Comm or any other software tool.
7. Seek for a reference in which expression (2.179) is derived in detail. Study its derivation and repeat it including your comments on intermediate derivations or simplifications. Save your work for a future reference.
8. Do a research on the techniques used in modern digital spectrum analyzers to estimate the frequency content of signals.
9. Study one of the techniques for implementing a 1-bit analog-to-digital converter and implement it using VisSim/Comm. Perform an analysis of the quantization error produced by the technique you have chosen and write down your conclusions, preferably comparing the results with those obtained with a conventional analog-to-digital conversion.

References

1. Benedetto, S. and Biglieri, E. Principles of Digital Transmission with Wireless Applications. New York, USA: Kluwer Academic and Plenum Publishers, 1999.
2. Breems, L. and Huijsing, J. H. Continuous-Time Sigma-Delta Modulation for A/D Conversion in Radio Receivers. New York, USA: Kluwer Academic Publishers, 2002.
3. Burns, P. Software Defined Radio for 3G. Norwood, USA: Artech House, Inc., 2003.
4. Chu, E. and George, A. Inside the FFT Black Box – Serial and Parallel Fast Fourier Transform Algorithms. Boca Raton, USA: CRC Press, 2000.
5. Cooley, J. W., and Tukey, J. W. An Algorithm for the Machine Calculation of Complex Fourier Series. Math. Comput. 19: pp. 297–301, 1965.
6. Couch, L.W. Digital and Analog Communication Systems, 7th Ed. Upper Saddle River, NJ, USA: Prentice Hall, Inc., 2007.
7. Girod, B., Rabenstein, R. and Stenger, A., Signals and Systems, Chichester, England: John Wiley and Sons, Inc., 2001.
8. Graf, R. F. Modern Dictionary of Electronics, 7th Ed, USA: Newnes & Butterworth-Heineman, 1999.
9. Hardy, G. H. and Rogosinski, W. W. Fourier Series. New York, USA: Dover Publications, 1999.
10. Haykin, S. Communication Systems, 3rd Ed. New York, USA: John Wiley and Sons, Inc., 1994.
11. Haykin, S. and Veen, B. V. Signals and Systems. New York, USA: John Wiley and Sons, Inc., 1999.
12. Howard, R. M. Principles of Random Signal Analysis and Low Noise Design - The Power Spectral Density and its Applications. New York, USA: John Wiley and Sons, Inc., 2002.
13. Hsu, H. P. (Theory and Problems of) Signals and Systems. Schaum's Outline Series. New York, USA: McGraw-Hill, Inc., 1995.
14. Jeruchim, M. C., Balaban, P. and Shanmugan, K. S. Simulation of Communication Systems – Modeling, Methodology, and Techniques. 2nd Ed. New York, USA: Kluwer Academic & Plenum Publishers, 2000.
15. Lathi, B. P. An Introduction to Random Signals and Communication Theory. USA: International Textbook Company, 1968.
16. Lathi, B. P. Modern Digital and Analog Communication Systems. 3rd Ed. New York, USA: Oxford University Press, 1998.
17. Lathi, B. P. Signal Processing and Linear Systems. Carmichael, CA, USA: Berkley-Cambridge Press, 1998.
18. Leon-Garcia, A. Probability, Statistics, and Random Processes for Electrical Engineering. 3rd Ed. Upper Saddle River, NJ, USA: Prentice Hall, Inc., 2008.
19. Lewis, A. D. A Mathematical Introduction to Signals and Systems – Time and Frequency Domain Representations of Signals. Math 334 course material. Queen's University, Canada (kindly provided by the author), September 2008.
20. Madisetti, V. K. and Williams, D. B. (Editors). The Digital Signal Processing Handbook. Atlanta, USA: CRC Press and IEEE Press, 1999.
21. Manolakis, D. G., Ingle, V. K. and Kogon, S. M. Statistical and Adaptive Signal Processing – Spectral Estimation, Signal Modeling, Adaptive Filtering, and Array Processing. USA: Artech House, Inc., 2005.
22. National Instruments – Digital Filter Design Toolkit for LabVIEW. Available at: http://www.ni.com/toolkits/df_design_toolkit.htm (last access on June 2009).
23. Oppenheim, A.V. and Schaffer, R.W., Discrete Time Signal Processing, Englewood Cliffs, NJ, USA: Prentice-Hall, Inc., 1989.
24. Oppenheim, A. V. Willsky, A. S. and Nawab, S. H. Signals and Systems, 2nd Ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

25. Park, S. Principles of Sigma-Delta Modulation for Analog-to-Digital Converters, Motorola Inc., APR8/D, Rev. 1, 1990.
26. Peebles, P. Z. Probability, Random Variables and Random Signal Principles. 2nd Ed. New York, USA: McGraw Hill, Inc., 1987.
27. Proakis, J. G. and Manolakis, D. G. Digital Signal Processing – Principles, Algorithms, and Applications, 3rd Ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.
28. Proakis, J. G. and Salehi, M. Communication Systems Engineering. 2nd Ed. Upper Saddle River, NJ, USA: Prentice Hall, Inc., 2002.
29. Proakis, J. G. Digital Communications. 3rd Ed. USA: McGraw Hill, Inc., 1995.
30. Proakis, J. G. (Editor). Wiley Encyclopedia of Telecommunications, Vols. 1–5. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2003.
31. Pursley, M. B. Introduction to Digital Communications. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2005.
32. Rorabaugh, C. B. Digital Filter Designer's Handbook: Featuring C++ Routines. USA: Tab Books / McGraw Hill, Inc., 1993.
33. Sheno, B. A. Introduction to Digital Signal Processing and Filter Design. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006.
34. Sklar, B. Digital Communications - Fundamentals and Applications. New Jersey, USA: Prentice Hall, Inc., 1988.
35. Smith, S. W. Scientist and Engineer's Guide to Digital Signal Processing, 2nd Ed. California Technical Publishing, 1999.
36. Stoica, P. and MOSES, R. L. Introduction to Spectral Analysis. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2005.
37. Titchworth, R. C. and WELCH, L. R. Power Spectra of Signals Modulated by Random and Pseudorandom Sequences. Pasadena, CA, USA: JPL Technical Report No. 32-140, October 1961.
38. Tuttlebee, W. (editor) Software Defined Radio – Origins, Drivers and International Perspectives. West Sussex, England: John Wiley & Sons, Inc., 2002.
39. Wikipedia – The Free Encyclopedia on line. Available at: <http://en.wikipedia.org> (last access on June 2009).
40. Weisstein, E. Wolfram MathWorld. Available at: <http://mathworld.wolfram.com> (last access on June 2009).
41. Xiong, F. Digital Modulation Techniques. Norwood, MA, USA: Artech House, Inc., 2000.

Digital Transmission

A Simulation-Aided Introduction with VisSim/Comm

Guimaraes, D.A.

2009, XXIII, 863 p. With CD-ROM., Hardcover

ISBN: 978-3-642-01358-4