

Digital Signal Processing

Yangang Cao

May 28, 2019

1 Introduction

1.1 Signals, Systems, and Signal Processing

Digital signal processing has developed rapidly over the past 40 years. A signal is defined as any physical quantity that varies with time, space, or any other independent variable or variables, a system may also be defined as a physical device that performs an operation on a signal, the method or set of rules for implementing the system by a program that performs the corresponding mathematical operations is called an algorithm. Signal processing is passing a signal through a system, this text deals with the processing of signals by digital means, either in software or in hardware.

Most of the signals encountered in science and engineering are analog in nature, but these kinds of signals can't be processed in computer. To perform the processing digitally, there is a need for an interface between the analog signal and the digital signal, this interface is called an *analog – to – digital(A/D) converter*. In applications where the digital output from the digital signal processor is to be given to the user in analog form, we must provide another interface from analog domain to the analog domain after processing, such an interface is called a *digital – to – analog(D/A) converter*.

Digital signal processing has some advantages compare to analog signal processing, such as flexible, accurate, easily stored and cheap, however, digital implementation has its limitations, one practical limitation is the speed of operation of A/D converts and digital signal processors.

1.2 Classification of Signals

1.2.1 Multichannel and Multidimensional Signals

If $s_k(t)$, $k = 1, 2, 3$, denotes the electrical signal from the k th sensor as a function of time, the set of $p = 3$ signals can be represented by a vector $S_3(t)$, where

$$S_3(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix}$$

We refer to such a vector of signals as a *multichannelsignal*.

If the signal is a function of M independent variables, this signal is called M – dimensional signal, such as

$$I(x, y, t) = \begin{bmatrix} I_r(x, y, t) \\ I_g(x, y, t) \\ I_b(x, y, t) \end{bmatrix}$$

1.2.2 Continuous-Time Versus Discrete-Time Signals

Continuous – time signals or *analog signals* are defined for every value of time and they take on value in the continuous interval (a, b) , these signals can be described by functions of a continuous variable. *Discrete – time signals* are defined only at certain specific values of time, these time instants need not be equidistant and can be represented mathematically by a sequence of real or complex numbers.

In applications, discrete -time signals may arise in two ways:

1. By selecting values of an analog signal at discrete-time instants.
2. By accumulating a variable over a period of time.

1.2.3 Continuous-Valued Versus Discrete -Values Signals

The values of a continuous or discrete-time signal can be continuous or discrete. If a signal takes on all possible values on a finite or an infinite range, it is said to be a *continuous – valued signal*. Alternatively, if the signal takes on values from a finite set of possible values, it is said to be a *discrete – valued signal*. A discrete-time signal having a set of discrete values is called a *digital signal*. The processing of converting a continuous-valued signal into a discrete-valued signal, called quantization, is basically an approximation process.

1.2.4 Deterministic Versus Random Signals

The mathematical analysis and processing of signals requires the availability of a mathematical description for the signal itself. This mathematical description, often referred to as the *signal model*, leads to another important classification of signals. Any signal that can be uniquely described by an explicit mathematical expression, a table of data, or a well-defined rule is called *deterministic*.

In many practical applications, however, there are signals that either cannot be described to any reasonable degree of accuracy by explicit mathematical formulas, or such a description is too complicated to be of any practical use. The lack of such a relationship implies that such signals evolve in time in an unpredictable manner. We refer to these signals as *random*.

It should be emphasized at this point that the classification of a *real – world* signal as deterministic or random is not always clear.

1.3 The Concept of Frequency in Continuous-Time and Discrete-Time Signals

We expect that the nature of time (continuous or discrete) would affect the nature of the frequency accordingly.

1.3.1 Continuous-Time Sinusoidal Signals

A simple harmonic oscillation:

$$x_a(t) = A \cos(\Omega t + \theta), -\infty < t < \infty$$

A is the amplitude of sinuoid, Ω is the frequency, and θ is the phase in radians. Instead of Ω , we often use the frequency F in cycles per second or hertz(Hz), where

$$\Omega = 2\pi F$$

The analog sinusoidal signal is characterized by the following properties:

- $x_a(t)$ is periodic:

$$x_a(t + T_p) = x_a(t)$$

where $T_p = 1/F$ is the fundamental period of the sinusoidal signal.

- Continuous-time sinusoidal signals with distinct frequencies are themselves distinct.
- Increasing the frequency F results in an increase in the rate of oscillation of the signal.

The relationships we have described for sinusoidal signals carry over to the class of complex exponential signals

$$x_a(t) = Ae^{j(\Omega t + \theta)}$$

This can easily be seen by expressing these signals in terms of sinusoids using the Euler identity

$$e^{\pm j\phi} = \cos \phi \pm j \sin \phi$$

A sinusoidal signal can be obtained by adding two equal-amplitude complex-conjugate exponential signals. Since a *positive frequency* corresponds to counter clockwise uniform angular motion, a *negative frequency* simply corresponds to clockwise angular motion.

1.3.2 Discrete-Time Sinusoidal Signals

A discrete-time sinusoidal signal:

$$x(n) = A \cos(\omega n + \theta), -\infty < n < \infty$$

where n is an integer variable, called the sample number, A is the amplitude of the sinusoid, ω is the frequency in radians per sample, and θ is the phase in radians.

The discrete-time sinusoids are characterized by following properties:

- A discrete-time sinusoid is periodic only if its frequency f is a rational number.
- Discrete-time sinusoids whose frequencies are separated by an integer multiple of 2π are identical.
- The highest rate of oscillation in a discrete-time sinusoid is attained when $\omega = \pi$ (or $\omega = -\pi$) or, equivalently, $f = \frac{1}{2}$ (or $f = -\frac{1}{2}$).

Usually, we choose the range $0 \leq \omega \leq 2\pi$ or $-\pi \leq \omega \leq \pi$ ($0 \leq f \leq 1, -\frac{1}{2} \leq f \leq \frac{1}{2}$), which we call the *fundamental range*.

1.3.3 Harmonically Related Complex Exponentials

These are sets of periodic complex exponentials with fundamental frequencies that are multiples of a single positive frequency. We can construct a linear combination of harmonically related complex exponentials by adding continuous-time, harmonically related exponentials, which is called *Fourier series expansion*. For discrete-time, $f_0 = 1/N$, they are only N distinct periodic complex exponentials, this is called *Fourier series*.

1.4 Analog-to-Digital and Digital-to-Analog Conversion

We view A/D conversion as a three-step process: sampling, quantization and coding. The accuracy of D/A conversion depends on the quality of D/A conversion, a simple form of D/A conversion is called zero-order hold or staircase approximation. The sampling rate is sufficiently high to avoid the problem commonly called *aliasing*.

1.4.1 Sampling of Analog Signals

We limit our discussion to *periodic* or *uniform sampling*, $F_s = 1/T$ is called the *sampling rate*(hertz). F is frequency of analog signal, the frequency variables F and f are linearly related as

$$f = \frac{F}{F_s}$$

and f is also called *relative* or *normalized frequency*.

We observe that the fundamental difference between continuous-time and discrete-time signals is in their range of values of the frequency variable F and f , or Ω and ω .

$$F_{max} = \frac{F_s}{2} = \frac{1}{2T}$$
$$\Omega_{max} = \pi F_s = \frac{\pi}{T}$$

For example, $F_2 = 50Hz$ is an alias of the frequency $F_1 = 50Hz$ when $F_s = 40Hz$. The relationship is

$$F_k = F_0 + kF_s, \quad k = \pm 1, \pm 2, \dots$$

$F_s/2$ is called *folding frequency*.

1.4.2 The sampling Theorem

F_s is selected that $F_s > 2F_{max}$, where F_{max} is the largest frequency component in the analog signal. The condition $F_s > 2F_{max}$ ensures that all the sinusoidal components in the analog signal are mapped into corresponding discrete-time frequency components with frequencies in the fundamental interval. The interpolation function is

$$g(t) = \frac{\sin 2\pi Bt}{2\pi Bt}$$

The sampling rate $F_N = 2B = 2F_{max}$ is called the *Nyquist rate*. Usually we sample the analog signal at a rate higher than Nyquist rate.

1.4.3 Quantization of Continuous-Amplitude Signals

The process of converting a discrete-time continuous-amplitude signal into a digital signal by expressing each sample value as a finite number of digits is called *quantization*, truncation and rounding are two ways of quantization, the quantization error is a sequence $e_q(n)$ defined as the difference between the quantized value and the actual sample value. The values allowed in the digital signal are called the *quantization levels*, whereas the distance δ between two successive quantization levels is called the *quantization step size* or *resolution*. The instantaneous quantization error cannot exceed half of the quantization step, in practice we can reduce the quantization error to an insignificant amount by choosing a sufficient number of quantization levels, quantization is an irreversible or noninvertible process.

1.4.4 Quantization of Sinusoidal Signals

Sinusoids are used as test signals in A/D converters. If the sampling rate F_s satisfies the sampling theorem, quantization is the only error in the A/D conversion process. The quality of the output of the A/D converter is usually measured by the *signal-to-quantization noise ratio (SQNR)*

$$SQNR = \frac{P_x}{P_q} = \frac{3}{2} \cdot 2^{2b}$$

Expressed in decibels (dB), the SQNR is

$$SQNR(\text{dB}) = 10 \log_{10} SQNR = 1.76 + 6.02b$$

1.4.5 Coding of Quantized Samples

The number of bits required in the coder is the smallest integer greater than or equal to $\log_2 L$.

2 Discrete-Time Signals and Systems

The major emphasis in this chapter is the characterization of discrete-time system in general and the class of linear time-invariant (LTI) systems in particular. Our motivation for the emphasis on the study of LTI system is twofold. First, there is a large collection of mathematical techniques that can be applied to the analysis of LTI systems. Second, many practical systems are either LTI systems or can be approximated by LTI systems.

2.1 Discrete-Time Signals

Discrete-time signal is not defined at instants between two successive samples, there are some alternative representations, such as functional, tabular and sequence.

2.1.1 Some Elementary Discrete-Time Signals

There are some basic signals

- Unit sample sequence
- Unit step signal
- Unit ramp signal
- Exponential signal

If the signal $x(n)$ is

$$x(n) = r^n e^{j\theta n}$$

its amplitude is

$$|x(n)| = A(n) = r^n$$

and phase is

$$\angle x(n) = \phi(n) = \theta n$$

2.1.2 Classification of Discrete-Time Signal

The energy E of a signal $x(n)$ is defined as

$$E = \sum_{-\infty}^{\infty} |x(n)|^2$$

If E is finite, then $x(n)$ is called an energy signal.

Many signals that possess infinite energy have a finite average power. The average power of $x(n)$ is defined as

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

Clearly, if E is finite, $P = 0$. On the other hand, if E is infinite, the average power P may be either finite or infinite. If P is finite, the signal is called *power signal*. Similarly, it can be shown that the complex exponential sequence $x(n) = Ae^{j\omega_0 n}$ has average power A^2 , so it is a power signal. On the other hand, the unit ramp sequence is neither a power signal nor an energy signal.

A signal $x(n)$ is *periodic* with period N ($N > 0$) if and only if

$$x(n+N) = x(n) \text{ for all } n$$

If there is no value of N , the signal is called *nonperiodic*. Periodic signals are power signals.

A real-valued signal $x(n)$ is called even if

$$x(-n) = x(n)$$

On the other hand, a signal $x(n)$ is called odd if

$$x(-n) = -x(n)$$

Clearly, we have

$$x_e(n) = \frac{1}{2}[x(n) + x(-n)]$$

$$x_o(n) = \frac{1}{2}[x(n) - x(-n)]$$

$$x(n) = x_e(n) + x_o(n)$$

2.1.3 Simple Manipulations of Discrete-Time Signals

A signal $x(n)$ may be shifted in time by replacing the independent variable n by $n - k$, where k is an integer. If k is a positive integer, the time shift results in a delay of the signal by k units of time. If the signal is not stored but is being generated by some physical phenomenon in real time, it is not possible to advance the signal in time, in real-time signal processing applications, the operation of advancing the time base of the signal is physically unrealizable.