

英文版

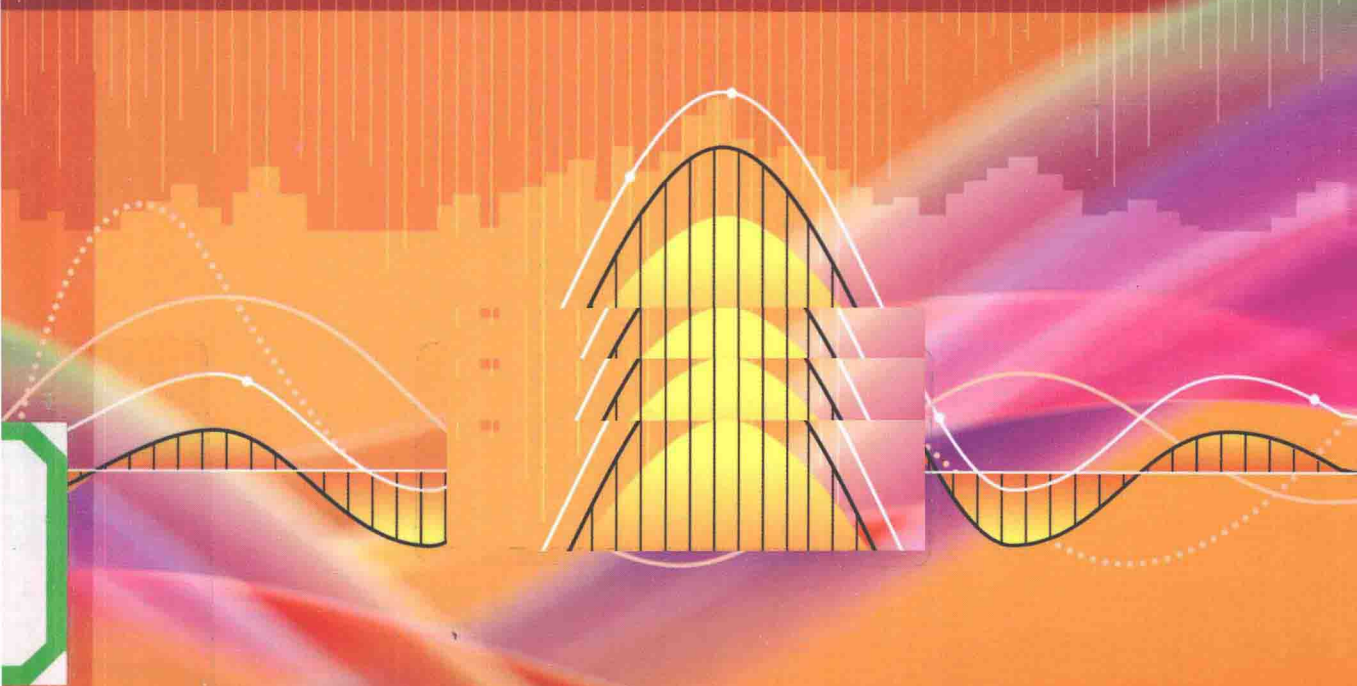


数字信号处理

(第2版)

Digital Signal Processing
Second Edition

© 蔡坤宝 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

英文版

数字信号处理

(第2版)

Digital Signal Processing
Second Edition

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书系统地阐述了数字信号处理所涉及的信号与系统分析和系统设计的基本理论、基本分析与设计方法、基本算法和处理技术。全书共 10 章, 主要内容包括: 离散时间信号与系统的基本概念, 离散时间信号与系统的变换域分析, 包括 z 变换和离散时间傅里叶变换、连续时间信号的抽样与重建, 离散傅里叶变换及其快速算法 (FFT), 数字滤波器实现的基本结构, IIR 和 FIR 数字滤波器的设计原理与基本设计方法, 数字信号处理中的有限长效应, 多抽样率数字信号处理。本书配有多媒体电子课件、英文版教学大纲、习题指导与实验手册。

本书可以作为电子与通信相关专业的本科数字信号处理课程中英文双语教学的教材, 或中文授课的英文版教学参考书, 也可供从事数字信号处理的工程技术人员学习参考。本书尤其适合初步开展数字信号处理课程中英文双语授课的师生选用。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

图书在版编目 (CIP) 数据

数字信号处理 = Digital Signal Processing: 英文 / 蔡坤宝编著. —2 版. —北京: 电子工业出版社, 2011.3
ISBN 978-7-121-12994-0

I. ①数… II. ①蔡… III. ①数字信号—信号处理—英文 IV. ①TN911.72

中国版本图书馆 CIP 数据核字 (2011) 第 029195 号

策划编辑: 王羽佳

责任编辑: 王羽佳

印 刷: 北京东光印刷厂

装 订: 三河市皇庄路通装订厂

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1092 1/16 印张: 20.75 字数: 832 千字

印 次: 2011 年 3 月第 1 次印刷

印 数: 4000 册 定价: 39.90 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前 言

21 世纪是国际化的知识经济时代,理工科教育已发生深刻的变化,各专业涉及的新理论与新技术发展日新月异,科技的创新在很大程度上已依赖于信息的及时获取、准确理解和有效利用。当今,数字信号处理理论和算法的研究、应用与实现技术的发展,以及其在现代信息与通信技术中的重要性和巨大潜力,已超越了初期人们所做的估计与预测。与此同时,社会对高素质信息技术人才的需求对高等学校专业基础课程的教学质量提出了越来越高的要求,然而“数字信号处理”及其相关技术的基础课程所能获得的学时数反而在减少。有知名专家与学者将当今的教学改革难题归结为:人类知识的无限积累与个人学习能力和时间的有限形成了日益尖锐的矛盾。

高等教育的国际化是当今教育与教学改革的必然趋势,国际视野和国际交流能力已成为我国高等学校人才培养的一项基本要求。为适应教学改革的新要求,我总结了“数字信号处理”教学与科研工作 20 余年所积累的经验与创新成果,并力图继承国内老一代专家与学者编著的优秀教材的知识体系结构严谨、系统性强的特色与传统,参考了 20 余本国内外一流或知名高等学校的优秀教材,通过消化、吸收和创新,编写了本书。2007 年 8 月本书第 1 版由电子工业出版社出版。此后,本书在作者本人执教的重庆大学通信工程学院本科数字信号处理课程双语教学班连续使用了 4 届。本书第 1 版于 2010 年已脱销,由于作者工作繁忙,时至今日才进行修订工作。

我在教材内容的选择、知识体系的组织和编排方面,做了慎重考虑——本书内容要适应我国高等学校的教学和课程设置的实际情况。面对数字信号处理知识内容迅速扩展和学时数有限的实际情况,我在编写过程中始终贯彻的基本思想是:使读者系统地掌握离散时间信号与系统分析与设计的基本理论;在两种常用的数字信号处理技术方面(基于 DFT 的连续时间信号的频谱分析、IIR 和 FIR 滤波器之类的数字信号处理系统的设计),力求使读者对分析与设计的原理和方法有较透彻的理解与掌握;在数字信号处理系统中的有限字长效应和多抽样率数字信号处理方面打下一定的基础;通过进一步自学或学习更加深入的后续课程,即可较容易地扩充数字信号处理的理论知识与实际技能。

基于我使用第 1 版作为教材的实际经验与体会,第 2 版保留了第 1 版中的主要内容,以适应目前本科教学的基本需要;压缩了篇幅,以适应学时数减少的实际情况;修正了第 1 版中的文字与公式符号错误,润色了语句文字。具体修订情况如下:

① 基于提高课堂教学效率和提高学生分析与解决问题能力的考虑,对第 1 版第 2、3 章中一些相对较简单的例题进行了精简。这些例题的题目被插入到相应章的习题中,可以作为学生课后作业。

② 考虑到学时数有限的实际情况,而且第1版未介绍频率抽样滤波器设计的内容,删除原6.3.4节。

③ 基于方便教师检验课堂教学效果的考虑,删除第1版附录F课后习题参考答案。为方便学生自学,课后习题参考答案可登录华信教育资源网(<http://www.hxedu.com.cn>)注册下载。

本书为使用本书作为教材的教师提供免费多媒体电子课件、英文版教学大纲、习题详解手册和实验指导手册等教学辅助资料,请教师登录华信教育资源网注册下载申请表格,填写后发邮件至 wyj@phei.com.cn 索取光盘。我认真准备了这套教学资源,相信对教学质量的提高大有裨益,可使初步开设数字信号处理双语教学课程的教师显著地提高备课效率。

在本书出版之际,谨向我的导师——重庆大学电气工程学院的江泽佳教授和周守昌教授表示由衷的感激和深深的敬意!衷心感谢为本书第1版的初稿提出评审意见及修改、校正的建设性建议的重庆大学电气工程学院的江泽佳教授、北京邮电大学电子工程学院的尹霄丽教授和英国伦敦 Bristol 大学的江丕书博士。衷心感谢为本书顺利出版而付出辛勤劳动的电子工业出版社高等教育分社谭海平社长(第1版责任编辑)、马岚副编审和王羽佳编辑。我还要向在数字信号处理双语教学班上使用了本书第1版作为教材的空军雷达学院等多所高校表示感谢!感谢他们对本书的信任与支持!

由于作者水平有限,本书的错误和疏漏之处,望读者批评指正!

蔡坤宝

2011年2月

Contents

1	Introduction	1
1.1	What Is a Signal?	1
1.2	What Is a System?	1
1.3	What Is Signal Processing?	2
1.4	Classification of Signals	2
1.4.1	Deterministic and Random Signals	2
1.4.2	Continuous-Time and Discrete-Time Signals	3
1.4.3	Periodic Signals and Nonperiodic Signals	4
1.4.4	Energy Signals and Power Signals	4
1.5	Overview of Digital Signal Processing	6
2	Discrete-Time Signals and Systems	7
2.1	Discrete-Time Signals: Sequences	7
2.2	Basic Sequences	10
2.2.1	Some Basic Sequences	10
2.2.2	Periodicity of Sequences	13
2.2.3	Representation of Arbitrary Sequences	15
2.3	Discrete-Time systems	16
2.4	Time-Domain Representations of LTI Systems	21
2.4.1	The Linear Convolution Sum	21
2.4.2	Interconnections of LTI Systems	24
2.4.3	Stability Condition of LTI systems	25
2.4.4	Causality Condition of LTI systems	26
2.4.5	Causal and Anticausal Sequences	26
2.5	Linear Constant-Coefficient Difference Equations	27
2.5.1	Recursive Solution of Difference Equations	27
2.5.2	Classical Solution of Difference Equations	28
2.5.3	Zero-Input Response and Zero-State Response	30
2.5.4	The Impulse Response of Causal LTI Systems	31
2.5.5	Recursive Solution of Impulse Responses	31
2.5.6	Classification of LTI Discrete-Time Systems	33
	Problems	34
3	Transform-Domain Analysis of Discrete-Time Signals and Systems	37
3.1	The z -Transform	37
3.1.1	Definition of the z -Transform	37
3.1.2	A General Shape of the Region of Convergence	37
3.1.3	Uniqueness of the z -Transform	40
3.2	Relation Between the ROCs and Sequence Types	41
3.3	The z -Transform of Basic Sequences	44
3.4	The Inverse z -Transform	45
3.4.1	Contour Integral Method	45
3.4.2	Partial Fraction Expansion Method	48
3.4.3	Long Division Method	50

3.4.4	Power Series Expansion Method	52
3.5	Properties of the z -Transform	53
3.6	The Discrete-Time Fourier Transform	60
3.6.1	Definition of the Discrete-Time Fourier Transform	60
3.6.2	Convergence Criteria	62
3.6.3	Properties of the Discrete-Time Fourier Transform	66
3.6.4	Symmetry Properties of the Discrete-Time Fourier Transform	68
3.7	Transform-Domain Analysis of LTI Discrete-Time Systems	70
3.7.1	The Frequency Response of Systems	71
3.7.2	The Transfer Function of LTI Systems	74
3.7.3	Geometric Evaluation of the Frequency Response	77
3.8	Sampling of Continuous-Time Signals	79
3.8.1	Periodic Sampling	79
3.8.2	Reconstruction of Bandlimited Signals	83
3.9	Relations of the z -Transform to the Laplace Transform	85
	Problems	88
4	The Discrete Fourier Transform	92
4.1	The Discrete Fourier Series	92
4.2	Properties of the Discrete Fourier Series	96
4.3	The Discrete Fourier Transform	100
4.4	Properties of the Discrete Fourier Transform	102
4.5	Linear Convolutions Evaluated by the Circular Convolution	109
4.6	Linear Time-Invariant Systems Implemented by the DFT	112
4.7	Sampling and Reconstruction in the z -Domain	114
4.8	Fourier Analysis of Continuous-Time Signals Using the DFT	117
4.8.1	Fourier Analysis of Nonperiodic Continuous-Time Signals	118
4.8.2	Practical Considerations	120
4.8.3	Spectral Analysis of Sinusoidal Signals	124
	Problems	126
5	Fast Fourier Transform Algorithms	130
5.1	Direct Computation and Efficiency Improvement of the DFT	130
5.2	Decimation-in-Time FFT Algorithm with Radix-2	131
5.2.1	Butterfly-Branch Transmittance of the Decimation-in-Time FFT	135
5.2.2	In-Place Computations	135
5.3	Decimation-in-Frequency FFT Algorithm with Radix-2	137
5.4	Computational Method of the Inverse FFT	139
	Problems	139
6	Digital Filter Structures	141
6.1	Description of the Digital Filter Structures	141
6.2	Basic Structures for IIR Digital Filters	142
6.2.1	Direct Form I	142
6.2.2	Direct Form II	143
6.2.3	Cascade Form	143
6.2.4	Parallel Form	145
6.3	Basic Structures for FIR Digital Filters	147
6.3.1	Direct Forms	147
6.3.2	Cascade Forms	148
6.3.3	Linear-Phase Forms	148
	Problems	150
7	Design Techniques of Digital IIR Filters	152
7.1	Preliminary Considerations	152

7.2	Frequency Response of Digital Filters	154
7.3	Discrete-Time Systems Characterized by Phase Properties	156
7.4	Allpass Systems	158
7.4.1	Nonminimum-Phase Systems Represented by a Cascade Connection	160
7.4.2	Group Delay of the Minimum-Phase Systems	161
7.4.3	Energy Delay of the Minimum-Phase Systems	162
7.5	Analog-to-Digital Filter Transformations	163
7.5.1	Impulse Invariance Transformation	164
7.5.2	Step Invariance Transformation	167
7.5.3	Bilinear Transformation	170
7.6	Design of Analog Prototype Filters	175
7.6.1	Analog Butterworth Lowpass Filters	175
7.6.2	Analog Chebyshev Lowpass Filters	179
7.7	Design of Lowpass IIR Digital Filters	184
7.7.1	Design of Lowpass Digital Filters Using the Impulse Invariance	184
7.7.2	Design of Lowpass Digital Filters Using the Bilinear Transformation	188
7.8	Design of IIR Digital Filters Using Analog Frequency Transformations	192
7.8.1	Design of Bandpass IIR Digital Filters	192
7.8.2	Design of Bandstop IIR Digital Filters	197
7.8.3	Design of Highpass IIR Digital Filters	202
7.9	Design of IIR Digital Filters Using Digital Frequency Transformations	206
7.9.1	Lowpass-to-Lowpass Transformation	207
7.9.2	Lowpass-to-Highpass Transformation	209
7.9.3	Lowpass-to-Bandpass Transformation	211
7.9.4	Lowpass-to-Bandstop Transformation	214
	Problems	216
8	Design of FIR Digital Filters	217
8.1	Properties of Linear Phase FIR Filters	217
8.1.1	The Impulse Response of Linear-Phase FIR Filters	218
8.1.2	The Frequency Response of Linear-Phase FIR Filters	220
8.1.3	Characteristics of Amplitude Functions	222
8.1.4	Constraints on Zero Locations	227
8.2	Design of Linear-Phase FIR Filters Using Windows	228
8.2.1	Basic Techniques	228
8.2.2	Window Functions	230
8.2.3	Design of Linear-Phase FIR Lowpass Filters Using Windows	236
8.2.4	Design of Linear-Phase FIR Bandpass Filters Using Windows	239
8.2.5	Design of Linear-Phase FIR Highpass Filters Using Windows	241
8.2.6	Design of Linear-Phase FIR Bandstop Filters Using Windows	242
	Problems	244
9	Finite-Wordlength Effects in Digital Signal Processing	246
9.1	Binary Number Representation with its Quantization Errors	246
9.1.1	Fixed-Point Binary Representation of Numbers	246
9.1.2	Floating-Point Representation	249
9.1.3	Errors from Truncation and Rounding	249
9.1.4	Statistical Model of the Quantization Errors	253
9.2	Analysis of the Quantization Errors in A/D Conversion	254
9.2.1	Statistical Model of the Quantization Errors	254
9.2.2	Transmission of the Quantization Noise through LTI Systems	257
9.3	Coefficient Quantization Effects in Digital Filters	258
9.3.1	Coefficient Quantization Effects in IIR Digital Filters	258
9.3.2	Statistical Analysis of Coefficient Quantization Effects	263
9.3.3	Coefficient Quantization Effects in FIR Filters	266
9.4	Round-off Effects in Digital Filters	268
9.4.1	Round-off Effects in Fixed-Point Realizations of IIR Filters	268

9.4.2	Dynamic Range Scaling in Fixed-Point Implementations of IIR Filters	274
9.5	Limit-Cycle Oscillations in Realizations of IIR Digital Filters	278
9.5.1	Zero-Input Limit Cycle Oscillations	278
9.5.2	Limit Cycles Due to Overflow	281
9.6	Round-off Errors in FFT Algorithms	288
9.6.1	Round-off Errors in the Direct DFT Computation	288
9.6.2	Round-off Errors in Fixed-point FFT Realization	289
	Problems	293
10	Multirate Digital Signal Processing	296
10.1	Sampling Rate Changed by an Integer Factor	296
10.1.1	Downsampling with an Integer Factor M	296
10.1.2	Decimation by an Integer Factor M	299
10.1.3	Upsampling with an Integer Factor L	302
10.1.4	Interpolation by an Integer Factor L	303
10.2	Sampling Rate Conversion by a Rational Factor	305
10.3	Efficient Structures for Sampling Rate Conversion	307
10.3.1	Equivalent Cascade Structures	308
10.3.2	Polyphase Decompositions	309
10.3.3	Polyphase Realization of Decimation Filters	310
10.3.4	Polyphase Realization of Interpolation Filters	311
	Problems	312
Appendix A	Tables for the z -Transform	315
Appendix B	Table for Properties of the Discrete-Time Fourier Transform	317
Appendix C	Table for Properties of the Discrete Fourier Series	318
Appendix D	Table for Properties of the Discrete Fourier Transform	319
Appendix E	Table for the Normalized Butterworth Lowpass Filters	320
References		321

1

Introduction

1.1 What Is a Signal?

Signals, in one form or another, constitute a basic ingredient of our daily lives. For example, a common form of human communication takes place through the use of speech signals, which may be in a face-to-face conversation or over a telephone channel. Another common form of human communication is visual in nature, with the signals taking the form of images of people or objects around us. Indeed, there are so many signals encountered in our living environment that the list of signals is almost endless.

Generally speaking, signals are a carrying body to convey information, while the information is contents embodied in signals. However, signals, in a narrow sense, are mathematically defined as a function of one or more independent variables that conveys information on nature of a physical phenomenon. When the function depends on a single independent variable, the signal is said to be one-dimensional. For Example, speech and music signals represent air pressure as a function of time at a point in space. When the function depends on two or more independent variables, the signal is said to be multidimensional signal. For example, a black-and-white picture is a representation of light intensity as a function of two spatial coordinates; a video signal in television consists of a sequence of images, called frames, and is a function of three independent variables that are two spatial coordinates and time.

Generally, a signal is a function of independent variables such as time, distance, position, temperature, pressure and etc. It is a common convention that the independent variable of the mathematical representation of a single variable signal will be represented to as time in this textbook, although it may in fact not represent time.

1.2 What Is a System?

A system, in its most general form, is defined as a combination and interconnection of several components to perform a designed task. For examples, the human physiology system, ecological system, communication system, electric power system and global positioning system are all the real-world systems, in a wide sense. However, a system, in a narrow sense, is mathematically defined as a transformation or operator that maps an input signal into an output signal. Specifically, a discrete-time system can be denoted as

$$y(n) = T[x(n)] \quad (1.1)$$

where $x(n)$ is the input signal, $y(n)$ is the output signal and 'T' is an operator which represents a rule or computation applied to the input signal to yield the output signal. Such a system is often depicted using a block diagram shown in Figure 1.1.

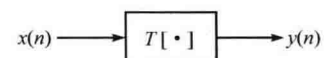


Figure 1.1 Block diagram representation for the system

A basic structure of commonly used communication systems is depicted in Figure 1.2. There are three basic elements in this system, namely, transmitter, channel and receiver. Functionally, the transmitter changes the message signal into a form suitable for transmission over the channel, the channel is the physical medium that connects the transmitter and receiver, and the receiver processes the channel output to produce an estimate of the message signal for a user.

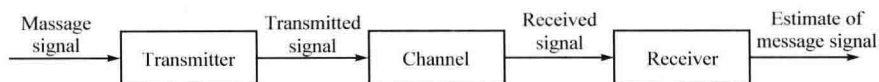


Figure 1.2 Basic structure of a communication system

1.3 What Is Signal Processing?

Signal processing is concerned with the representation, transformation, and manipulation of signals and the information they contain. For example, we may wish to separate two or more signals that have been mixed together, or we may want to enhance some signal components or estimate some parameters of a signal model. In communication systems, it is commonly necessary to do pre-processing such as modulation, signal conditioning, and compression prior to transmission over a channel and then to post process at the receiver. The technology for signal processing was almost exclusively continuous-time analog technology until 1960s. The rapid development of digital computers and micro processors together with some important theoretical progress such as the fast Fourier transform (FFT) algorithm caused a major shift to digital technologies, giving rise to the field of digital signal processing.

1.4 Classification of Signals

In this textbook, we will restrict our attention to one-dimensional signals which are defined as single-valued functions of independent variable time. “Single-valued” means that for every specified instant of time there is a unique value of the function except for the discontinuities of the function. The value of a signal at a specified time is called its amplitude. The variation of the amplitude as a function of the independent time variable is called its waveform. The classification of signals is a basic problem in the field of signal processing, because different types of signals concern with different representations and processing methods.

1.4.1 Deterministic and Random Signals

According to the certainty of some features of general signals, the signals can be classified into two sets, that is, deterministic signals and random signals.

1. Deterministic Signals

A deterministic signal is such a signal about which there is no uncertainty with respect to its value at any specified time. Thus, a deterministic signal can be completely described by a known function of time.

A typical deterministic signal is a well-known sinusoidal signal, that is,

$$x_c(t) = A \sin(\Omega_0 t + \theta) \quad (1.2)$$

where A is its amplitude, Ω_0 is its angular frequency with units radians per second (rad/s), and θ is its initial phase with units radians (rad).

2. Random Signals

A random signal is such a signal about which there is uncertainty before it occurs. In other words, a random signal is generated in a random fashion and cannot be predicted ahead of time. Thus, a random signal cannot be described by a deterministic function. According to the theory of random processes, such a signal may be viewed as one realization of an ensemble of signals, with each signal in the ensemble having a different waveform. Moreover, each signal within the ensemble has a certain probability of occurrence. The ensemble of signals is called a random process.

A typical random signal is the random initial-phase sinusoidal signal, that is,

$$x_c(t) = A \sin(\Omega_0 t + \varphi) \quad (1.3)$$

where the amplitude A and the angular frequency Ω_0 are constants, and the initial phase φ is a random variable with a probability density function $p(\varphi) = 1/2\pi$. Although the amplitude and angular frequency of the signal are constants, the initial phase cannot be predicted before it is

generated.

Another random signal is thermal noise generated in the amplifier of a radio or television receiver. Its amplitude fluctuates between positive and negative values in a complete random fashion.

1.4.2 Continuous-Time and Discrete-Time Signals

According to the continuity of the independent time variable for signals, the signals can be classified into two classes as follows.

1. Continuous-Time Signals

A signal $x(t)$ is said to be a continuous-time signal, if it is defined in the continuous-time domain. However, it is not necessary for the amplitude of the signal to be continuous at any time instants. In other words, a continuous-time signal may be undefined at a finite number of discrete time instants.

Furthermore, the continuous-time signals can be classified into two subclasses. One is the analog signals whose time variable and amplitude are all continuous. Such a signal is shown in Figure 1.3(a). Another is the quantized boxcar or staircase signals whose time variable is continuous while its amplitude takes discrete values with finite precision. Such a signal is shown in Figure 1.3(b).

2. Discrete-Time Signals

A discrete-time signal is defined only at discrete time instants. Thus, the independent variable of the signal takes discrete values only, which are usually uniformly spaced on the time axis. However, the amplitude of a discrete-time signal may take infinite-precision values or finite-precision values. A discrete-time signal with discrete-valued amplitude represented by a finite number of digits is the so-called digital signal.

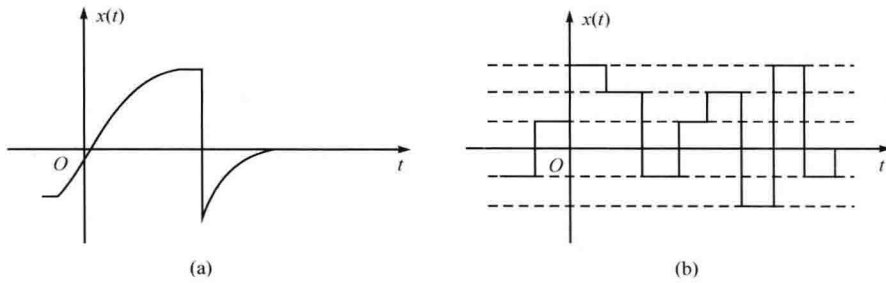


Figure 1.3 Continuous-time signals: (a) an analog signal; (b) a quantized boxcar signal

A discrete-time signal is often derived from a continuous-time signal by sampling it at a uniform rate, that is,

$$x(n) = x_c(nT) = x_c(t)|_{t=nT}, \quad n = 0, \pm 1, \pm 2, \dots \quad (1.4)$$

where T denotes the sampling period with units seconds (s), n denotes an integer that may assume positive and negative values, however, it corresponds to time. Such a signal is shown in Figure 1.4(a), where the amplitude values of the signal are assumed to be continuous, or infinitely precise. A corresponding digital signal may be obtained by taking the quantized amplitude values of the signal in Eq.(1.4), which is shown in Figure 1.4(b).

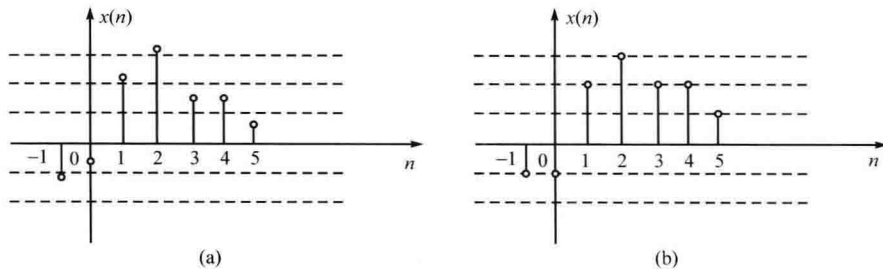


Figure 1.4 (a) A discrete-time signal; (b) a digital signal

1.4.3 Periodic Signals and Nonperiodic Signals

A continuous-time signal $x(t)$ is periodic if and only if it satisfies

$$x(t) = x(t + T_0), \text{ for all } t \quad (1.5)$$

where T_0 is a positive number called the period of the signal with units seconds (s). Specifically, the smallest value of T_0 satisfying Eq.(1.5) is referred to as the fundamental period of the periodic signal $x(t)$. Accordingly, the fundamental period T_0 defines the duration of one complete cycle of the periodic signal $x(t)$.

The reciprocal of the fundamental period T_0 is called the fundamental frequency of the periodic signal $x(t)$, which is denoted as

$$f_0 = 1/T_0 \quad (1.6)$$

with units hertz (Hz). It describes how frequently the periodic signal repeats itself. Therefore, the fundamental frequency is also measured in cycles per second.

The fundamental angular frequency of the periodic signal $x(t)$ with units radians per second (rad/s) is defined by

$$\Omega_0 = 2\pi f_0 = 2\pi / T_0 \quad (1.7)$$

Since there are 2π radians in one complete cycle, Ω_0 is often referred to simply as the frequency.

Any signal $x(t)$ for which no value of T_0 satisfies Eq.(1.5) is called a nonperiodic signal, or aperiodic signal.

A discrete-time signal $x(n)$ is said to be periodic if it satisfies

$$x(n) = x(n + N), \text{ for all } n \quad (1.8)$$

where N is a positive integer called the period of the signal $x(n)$. The smallest integer N for which Eq.(1.8) satisfies is called the fundamental period of the periodic signal $x(n)$. The fundamental angular frequency of the periodic signal $x(n)$, or simply, fundamental frequency with units radians is defined by

$$\omega_1 = 2\pi / N \quad (1.9)$$

1.4.4 Energy Signals and Power Signals

It is well-known that the instantaneous power dissipated in a resistor with 1 ohm resistance is represented as

$$p(t) = v^2(t) / R = v^2(t) \quad (1.10)$$

where $v(t)$ is assumed to be the voltage across the resistor with units volts (V), with the result that the instantaneous power $p(t)$ is measured in watts (W). Equivalently, we have

$$p(t) = i^2(t)R = i^2(t) \quad (1.11)$$

where $i(t)$ is the current flowing through the unit resistor (with 1 ohm resistance). In both cases, the instantaneous power $p(t)$ is equal to the square of the signal $v(t)$ or $i(t)$ for the unit resistor R . Thus, in the signal analysis, it is customary to define the power of a signal in terms of the unit resistor, regardless of whether a given signal $x(t)$ represents a voltage or a current.

Considering the above convention, we discuss the power and energy of different signals as follows.

1. Aperiodic Continuous-Time Signals

For an arbitrary continuous-time signal $x(t)$ which may be complex-valued, the instantaneous power normalized to unit resistance is defined as

$$p(t) = |x(t)|^2 \quad (1.12)$$

with units watts. The total energy of the continuous-time signal $x(t)$ is defined as

$$E = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} |x(t)|^2 dt = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (1.13)$$

with units joules (J). Meanwhile, the average power of the signal is defined as

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt \quad (1.14)$$

with units watts.

2. Periodic Continuous-Time Signals

If $x_p(t)$ is a continuous-time periodic signal with period T_0 , then the nontrivial energy of the whole signal is always infinite. However, the average power of the whole signal is equal to the average power in any period of the signal, that is,

$$P = \frac{1}{T_0} \int_{t_0}^{t_0+T_0} |x_p(t)|^2 dt \quad (1.15)$$

where t_0 is an arbitrary time constant. Particularly, for a real-valued signal $x_p(t)$, the square root of the average power P is called the root mean-square value of the periodic signal.

3. Aperiodic Discrete-Time Signals

In signal processing field, the total energy of a discrete-time signal $x(n)$, is conventionally defined as

$$E = \sum_{n=-\infty}^{\infty} |x(n)|^2 \quad (1.16)$$

It is measured in square volts (V^2). The average power of the signal $x(n)$ is defined as

$$P = \lim_{M \rightarrow \infty} \frac{1}{M+1} \sum_{n=-M/2}^{M/2} |x(n)|^2 \quad (1.17)$$

In the field of electrical engineering, the discrete-time signal $x(n)$ is generally obtained by sampling a continuous-time signal $x_c(t)$. The energy for such a discrete-time signal is meaningfully defined as

$$E = \lim_{M \rightarrow \infty} T \sum_{n=-M/2}^{M/2} |x_c(nT)|^2 = T \sum_{n=-\infty}^{\infty} |x(n)|^2 \quad (1.18)$$

where T is the sampling period measured in seconds, which is assumed to be small enough. The average power for such a signal is also computed by using the Eq.(1.17). However, its physically meaningful units are given by watts (W).

4. Periodic Discrete-Time Signals

For a periodic discrete-time signal $x_p(n)$, the nontrivial energy of the whole signal is infinite, while the average power of the signal is given by

$$P = \frac{1}{N} \sum_{n=n_0}^{N+n_0-1} |x_p(n)|^2 \quad (1.19)$$

where N is the period of the signal and n_0 is an arbitrary integer. For the same reason, if the signal $x_p(n)$ is derived from a continuous-time periodic signal $x_p(t)$, then the average power represented in Eq.(1.19) should be measured in watts (W).

In the field of signal processing, the classification of signals in terms of energy and power is very useful. A signal is referred to as an energy signal if and only if the total energy of the signal satisfies the condition

$$0 < E < \infty \quad (1.20)$$

so that the average power of the signal must be zero. A signal is referred to as a power signal if and only if the average power of the signal satisfies the condition

$$0 < P < \infty \quad (1.21)$$

this implies that the total energy of the whole signal must be infinite. Thus we conclude that the energy and power classifications of signals are mutually exclusive.

It should be noted that a signal satisfying neither inequality (1.20) nor inequality (1.21) is neither an energy signal nor a power signal.

1.5 Overview of Digital Signal Processing

In this modern world we are surrounded by all kinds of signals in various forms. In practical engineering, signals are carriers of information, both useful and unwanted. Therefore, extracting or enhancing the useful information from a mix of conflicting information is a simplest form of signal processing. Generally speaking, signal processing is an operation designed for extracting, enhancing, storing, and transmitting useful information.

The signals that we encounter in practice are mostly analog signals, which vary continuously in time and amplitude, are processed using electrical networks containing active and passive circuit elements. This approach is known as analog signal processing (ASP). This processing can be roughly described by the block diagram shown in Figure 1.5.

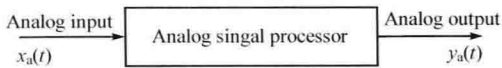


Figure 1.5 Block diagram for analog signal processing

The analog signal processing can be equivalently performed by digital signal processing (DSP), which uses such a structure as shown in Figure 1.6. The function of every block element in Figure 1.6 is introduced below.

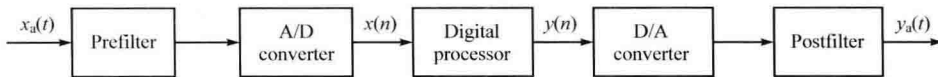


Figure 1.6 Equivalent analog signal processing

(1) The prefilter or antialiasing filter is used to condition the analog signal $x_a(t)$ to prevent aliasing.

(2) The A/D (analog-to-digital) converter produces a stream of binary numbers, denoted by a sequence $x(n)$.

(3) The digital processor processes the binary data in terms of a designed signal processing algorithm, which results in a stream of binary numbers, denoted by a sequence $y(n)$. The digital signal processor can represent a general-purpose computer, or digital hardware, and so on.

(4) The D/A (digital-to-analog) converter performs the inverse operation with respect to the A/D converter. It produces a staircase waveform from a sequence $y(n)$ of binary numbers, which is the first step toward producing a desired analog output signal $y_a(t)$.

(5) The postfilter is an analog lowpass filter used to smooth out staircase waveform into a desired analog signal $y_a(t)$.

Comparing above two approaches to process an analog signal, the DSP approach is more complicated than the ASP approach. Therefore, one might ask a question: Why process analog signals digitally? The answer lies on many advantages offered by the DSP.

A major drawback of ASP is its limited scope for performing complicated signal processing applications. This translates into nonflexibility in processing and complexity in system designs. All of those generally lead to expensive products. On the other hand, using a DSP approach, it is possible to convert an inexpensive personal computer into a powerful signal processor. Some important advantages of DSP approach may be as follows:

(1) Systems using the DSP approach can be developed using software running on a general-purpose computer. Therefore, DSP is relatively convenient to develop and test, and the software is portable.

(2) DSP operations are based essentially on additions and multiplications, leading to extremely stable processing capability.

(3) DSP operations can easily be modified in real time, often by simple programming changes, or by reloading registers.

(4) DSP has lower cost due to VLSI technology, which reduces costs of memories, gates, microprocessors, and so on.

The principal disadvantage of DSP is the speed of operations, especially at very high frequencies. Primarily due to the above advantages, DSP is now becoming a first choice in many technologies and applications, such as consumer electronics, communications, wireless telephones, and medical imaging.

2

Discrete-Time Signals and Systems

In this chapter, we first discuss the time-domain representation of a discrete-time signal as a sequence of numbers. We then describe some basic discrete-time signals or sequences that play important roles in the time-domain characterization of arbitrary discrete-time signals and discrete-time systems. A number of basic operations on discrete-time signals are discussed next. The discussion of discrete-time systems begins with the mathematical definition of discrete-time systems and the properties of such systems, including the linearity, time-invariance, stability and causality of the systems. We will concentrate our particular attention on the time-domain representation of linear time-invariant (LTI) systems through the linear convolution sum. Finally, we conclude this chapter with the discussion of a class of linear time-invariant systems represented by linear constant-coefficient difference equations and their solutions.

2.1 Discrete-Time Signals: Sequences

As mentioned earlier, a discrete-time signal $x(n)$ is represented as a sequence of numbers called samples, where n is the independent variable of the signal and takes integer numbers only. Therefore, a discrete-time signal is actually a set of numbers with an integer index n and can be mathematically denoted as

$$\{x(n)\}, \quad -\infty < n < \infty \quad (2.1)$$

Strictly speaking, $x(n)$ represents the n th element of the set, or the n th sample of the discrete-time signal. However, it is customary to use $x(n)$ to denote the sequence, if it does not cause confusion in concerned contexts. For example, a discrete-time signal can be formally expressed as

$$\{x(n)\} = \{(1/2)^n\}, \quad n \geq 0 \quad (2.2)$$

Of course, we would like to express this signal into

$$x(n) = \begin{cases} (1/2)^n, & n \geq 0 \\ 0, & n < 0 \end{cases} \quad (2.3)$$

Here the notation $x(n)$ represents the whole discrete-time signal, or the sequence representing the signal. Its n th sample for $n \geq 0$ has the value $(1/2)^n$. The graphical representation of this signal is illustrated in Figure 2.1, which is also called the waveform of the signal.

In practical engineering, a discrete-time signal $x(n)$ is often obtained by periodically sampling a continuous-time $x_c(t)$ at uniformly spaced time points, which leads to

$$x(n) = x_c(nT) = x_c(t)|_{t=nT}, \quad n = 0, \pm 1, \pm 2, \dots \quad (2.4)$$

where the time spacing T between two consecutive samples of the discrete-time signal is called the sampling interval or sampling period, with units seconds (s). The reciprocal of the sampling interval is

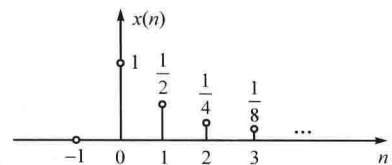


Figure 2.1 Graphical representation of a discrete-time signal

called the sampling frequency with units hertz (Hz), that is, $f_s = 1/T$ (Hz). Here, we have assumed that the sampling interval T is small enough to satisfy the recovery condition of original continuous-time signal $x_c(t)$ from the samples of the sequence $x(n)$, which is stated as the sampling theorem to be discussed in Chapter 4.

In the analysis of discrete-time signal processing systems, sequences are manipulated in several basic ways. We now briefly describe basic operations on sequences as follows.

1. Sample Addition

Let $x_1(n)$ and $x_2(n)$ denote a pair of sequences. The sum sequence $y(n)$ obtained by a sample-by-sample addition is defined by

$$y(n) = x_1(n) + x_2(n) \quad (2.5)$$

2. Sample Multiplication

Let $x_1(n)$ and $x_2(n)$ denote a pair of sequences. The product sequence $y(n)$ obtained by a sample-by-sample multiplication is defined by

$$y(n) = x_1(n)x_2(n) \quad (2.6)$$

3. Scalar Multiplication

If $x(n)$ is a sequence, then a scaled sequence $y(n)$ resulted from a scalar multiplication of $x(n)$ is defined by

$$y(n) = cx(n) \quad (2.7)$$

where c is an arbitrary constant. Clearly, it is a special case of the sample multiplication. Here, one out of two sequences has constant sample values, c .

4. Sample Accumulation

This differs from the sample addition operation. It adds all sample values of sequence $x(n)$, starting from $-\infty$ to n , that is,

$$y(n) = \sum_{k=-\infty}^n x(k) \quad (2.8)$$

where the resulting sequence $y(n)$ is called the accumulated sequence of $x(n)$.

Example 2.1 Let $x(n) = \begin{cases} (1/2)^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$. Determine its accumulated sequence $y(n)$.

Solution: From the definition in Eq.(2.8) and the given sequence in this problem, we know that if $n < 0$, then $y(n) = 0$, and if $n \geq 0$, then

$$y(n) = \sum_{k=-\infty}^n x(k) = \sum_{k=0}^n (1/2)^k = 2 - (1/2)^n, \quad \text{for } n \geq 0$$

Thus, the desired sequence is given by

$$y(n) = \begin{cases} 2 - (1/2)^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

5. Time Shifting

Let $x(n)$ denote a sequence. The time-shifted version of $x(n)$ is defined by

$$y(n) = x(n - n_0) \quad (2.9)$$

where n_0 is an integer. If n_0 is positive, then the sequence $y(n)$ is obtained by shifting each sample of $x(n)$ to the right with n_0 sample intervals. If n_0 is negative, then each sample of $x(n)$ is shifted to the left with $|n_0|$ sample intervals.