Unit 1

Text

Preface to Modern Engineering Mathematics

1.1 Modern Developments and the Teaching of Maths to Engineers

Developments in computer technology and related software have provided the engineer with tools of increasing power and sophistication which have significant implications for the use and role of mathematics in engineering practice[1]. In particular, it has led to greater use of mathematical modelling and simulation as the basis for the analysis and design of engineering systems, thus providing a more flexible and economic approach to the traditional methods which relied heavily on costly experimentation and the building of scaled models. Clearly such developments, particularly those in computer algebra or symbolic manipulation packages, also have important implications for the teaching of mathematics on engineering degree courses. To the inexperienced it is tempting to believe that the use of packaged software solves all the problems of analysis that an engineer is likely to meet and thus eliminates the need for engineering students to study mathematics. On the contrary, to the experienced engineer, the dangers of using packaged software as ‘black box’ solution generators are well understood yet cannot be overstressed. *If engineers are to take full advantage of sophisticated computational tools then it is essential that they become effective at mathematical modelling and discriminating, intelligent and wary users of packaged software and other aids to computational modelling*[2]. The need for mathematical skills is, therefore, greater than ever but it is widely recognized that, as a consequence of these computer related developments, there is a need for a shift in emphasis in the teaching of mathematics to students studying engineering. This shift is away from the simple mastery of solution techniques and towards development of a greater understanding of mathematical ideas and processes together with efficiency in applying this understanding to the formulation and analysis of mathematical models of physical phenomena and engineering systems. However, it is recognized that the development of understanding and the mastery of solution techniques are not mutually exclusive objectives. *There is little doubt that a high degree of fluency in the manipulation of mathematical expressions will always be required, for without this there can be no real understanding*[3]. The challenge to the teacher is that of achieving the correct balance in the mathematics curriculum.

1.2 Skills Development and Learning by Doing

The objective of the authoring team in writing this book is to achieve a balance between the development of understanding and the mastery of solution techniques with the emphasis being on the development of students’ ability to use mathematics with understanding to solve engineering problems.

Worked examples

Consequently, this book is not a collection of recipes and techniques designed to teach students to solve routine exercises, nor is mathematical rigour introduced for its own sake[4]. It contains over 350 worked examples, many of which incorporate mathematical models and are designed both to provide relevance and to reinforce the role of mathematics in various branches of engineering.

Applications

To provide further exposure to the use of mathematical practice, each chapter contains sections on engineering applications. These sections form an ideal framework for individual, or group, case study assignments leading to a written report and/or oral presentation; thereby helping to develop the skills of mathematical modelling necessary to prepare for the more open-ended modelling exercises at a later stage of the course.

Exercises

There are numerous exercise sections throughout the text and at the end of each chapter there is a comprehensive set of review exercises. While many of the exercise problems are designed to develop skills in mathematical techniques, others are designed to develop understanding and to encourage learning by doing, and some are of an open-ended nature. This book contains over 1000 exercises and answers to all the questions are given. It is hoped that this provision, together with the large number of worked examples and style of presentation, also makes the book suitable for private or directed study.

Numerical methods

Recognizing the increasing use of numerical methods in engineering practice, which often complement the use of analytical methods in analysis and design and are of ultimate relevance when solving complex engineering problems, there is wide agreement that they should be integrated within the mathematics curriculum. Consequently the treatment of numerical methods is integrated with the analytical work throughout the book. Algorithms are written in pseudocode and are, therefore, readily transferable to any specific programming language by the user.

1.3 Content

The range of material covered in the book is regarded as appropriate for a first level core studies course in mathematics for undergraduate courses in all engineering disciplines. The choice of material also reflects the proposals contained in the report *A Core Curriculum in Mathematics for the European Engineer* published by The European Society for Engineering Education (SEFI) in 1992. Whilst designed primarily for use by engineering students it is believed that the book is also highly suitable for students of the physical sciences and applied mathematics. Material appropriate for second level undergraduate core studies, or possibly elective studies for some engineering disciplines, is contained in the companion text *Advanced Modern Engineering Mathematics. As a* *result of the widening of access opportunities, particularly in the United Kingdom, there is increasing heterogeneity in background knowledge in mathematics of students entering degree courses in engineering*[5].

Chapter 1 deals with numbers, algebra and functions and includes sections on topics, such as trigonometric, exponential and logarithmic functions, a knowledge of which has traditionally been assumed on entry to an engineering degree course. To most students such sections will provide a review of material with which they already have some familiarity. The remainder of the chapter develops the material further and includes sections on computer arithmetic and numerical evaluation of functions.

Chapter 2 extends the number system to include complex numbers which have important applications in engineering. Vector and matrix techniques provide the framework for much of the developments in modern engineering and so the engineer needs to have a good understanding of the foundations of vector and linear algebra. Consequently **Chapters 3 and 4** are devoted to these topics with the material being further developed in the companion text. With the increasing importance of software engineering and use of expert systems, discrete mathematics is receiving prominence, with importance in many branches of engineering. While it may not be necessary for all engineers to have a deep understanding of discrete mathematics it is believed essential that they should all have a familiarity and ease with the relevant notations and formalism; this therefore is the main objective of **Chapter 5**.

Chapter 6 provides a basic introduction to the ideas of sequences, series and limits, as essential prerequisites for the study of the calculus, which remains a powerful mathematical tool for use in solving engineering problems.

Chapters 7 and 8 are devoted to the calculus of functions of one variable and, recognizing again the mixed background knowledge in mathematics of the students, the basic ideas and techniques of differentiation and integration are reviewed in Chapter 7.

Chapter 9 extends the calculus to the case of functions of more than one variable.

Chapters 10 and 11 relate to ordinary differential equations which are representative mathematical models of practical problems in various branches of engineering. Fourier series analysis is central to many applications in engineering, such as the analysis and design of oscillatory and nonlinear systems.

Chapter 12 provides a brief introduction to this topic appropriate for first level study, with a more detailed treatment being given in the companion text. Engineering is a discipline founded upon experiment, and engineers need to know how to process their experimental data and how to assess the results of others' experiments. The aim of statistics is to extract useful information from the data. The book concludes with **Chapter 13** which illustrates how data may be plotted to good effect, and then goes on to cover the essential probability theory necessary to take account of uncertainty in engineering.

1.4 Acknowledgements

It is a pleasure to acknowledge individuals who have contributed to the development of the book, in particular Nigel Steele of Coventry University, who is a member of the authoring team of the companion text, John Berry of the Polytechnic of the South West, for his contribution at the outset, and to the many reviewers for giving up their valuable time to contribute comments and suggestions during the preparation of the manuscript. The authoring team have been fortunate in having a superb production team at Addison-Wesley that has given every form of assistance throughout the preparation period. The team wish to thank all those concerned and in particular the development director, Sarah Mallen, for her continued enthusiasm and support and to Susan Keany for her diligence and patience as production editor.

Words & Expressions

Algebra	['ældʒibrə]	*n.*	代数学
algorithm	['ælgəriðəm]	*n.*	算法
arithmetic	[ə'riθmətik]	*n.*	算术
calculus	['kælkjuləs]	*n.*	微积分
curriculum	[kə'rikjuləm]	*n.*	课程
differentiation	[ˌdifəˌrenʃi'eiʃən]	*n.*	微分
discipline	['disəplin]	*n.*	学科
discrete	[dis'kriːt]	*adj.*	离散的
equation	[i'kweiʃən]	*n.*	方程式；等式
evaluation	[iˌvælju'eiʃən]	*n.*	赋值；值的计算
exponential	[ˌekspəu'nenʃəl]	*adj.*	指数的，幂数的
formalism	['fɔːməlizəm]	*n.*	形式
heterogeneity	[ˌhetərəudʒi'niːiti]	*n.*	异种，异质，不同成分
integration	[ˌinti'greiʃən]	*n.*	积分
limit	['limit]	*n.*	极限
linear	['liniə]	*adj.*	线性的（nonlinear，非线性的）
logarithmic	[ˌlɔgə'riθmik]	*adj.*	对数的
matrix	['meitriks]	*n.*	矩阵
oscillatory	['ɔsileitəri]	*adj.*	振荡的；变动的
sequence	['siːkwəns]	*n.*	序列
statistics	[stə'tistiks]	*n.*	统计；统计学
trigonometric	[trigənə'metrik]	*adj.*	三角学的；三角法的
variable	['vɛəriəbl]	*n.*	变量
vector	['vektə]	*n.*	矢量，向量
Expert system			专家系统
Fourier series			傅里叶级数
Oral presentation			口头报告
Ordinary Differential Equation			常微分方程（*abbr*. ODE）

Notes

[1] Developments in computer technology and related software have provided the engineer with tools of increasing power and sophistication which have significant implications for the use and role of mathematics in engineering practice. 本句中 which 引导的定语从句修饰整个主句，have implication for…表示“暗示”。

计算机技术及其相关软件的发展为工程技术人员提供了功能强大且日趋完善的工具，使数学应用于工程实践并发挥着重要作用。

[2] If engineers are to take full advantage of sophisticated computational tools then it is essential that they become effective at mathematical modelling and discriminating, intelligent and wary users of packaged software and other aids to computational modelling. 由 that 引导的主语从句（it 是形式主语）中 become 做谓语，后接两个并列成分 effective at…和…users of…，其中 discriminating, intelligent 及 wary 均修饰 users。

若想充分利用这些复杂的计算工具，工程技术人员必须熟悉数学建模，能区别各种软件包及其他计算建模辅助工具，并能灵活运用。

[3] There is little doubt that a high degree of fluency in the manipulation of mathematical expressions will always be required, for without this there can be no real understanding. 本句中 for 表示原因，manipulation 原意为“操纵、控制”，这里转译为“运用”。

毫无疑问，学生必须熟练运用数学公式，否则就不可能有真正意义上的理解。

[4] Consequently, this book is not a collection of recipes and techniques designed to teach students to solve routine exercises, nor is mathematical rigour introduced for its own sake. 本句中 not (neither)…nor…，表示“既不……也不……”，for one’s own sake 作“为了他自己的好处”解。

因此，本书既不是教授学生解题方法与技巧的汇总，也未以数学本身所要求的严谨方式编写。

[5] As a result of the widening of access opportunities, particularly in the United Kingdom, there is increasing heterogeneity in background knowledge in mathematics of students entering degree courses in engineering. 本句中 access 原意“通路，访问，入门”，heterogeneity 原意“异种，异质，不同成分”，翻译时难度较大，需要在充分理解的基础上意译。

由于入学机会增多，修读工程类学位课程学生的数学基础愈见参差不齐，尤以英国为甚。

Grammar

科技英语的语法特点

科技英语作为一种揭示客观外部世界的本质和规律的信息传递工具，具有准确、简明扼要和客观正式等特点。科技文章文体的特点是：语言简练、结构严谨、逻辑性强、原理概念清楚、重点突出、段落章节分明。具体而言，科技英语在用词、语法结构及表达方式上有其自身的特点，下面分别予以介绍。

1．词汇

① 大量使用专业词汇和半专业词汇，例如，calculus（微积分学）, bandwidth（带宽）, flip-flop

（触发器）等是专业词汇，而 series, work 等是半专业词汇，在不同的学科领域其含义有所不同，例如，series 可作“级数”（数学）解，也可作“串联”（电学）解。

② 大量使用词缀和词根，例如，外语教学与研究出版社出版的《英汉双解信息技术词典》中以 tele-构成的单词有 30 多个。

③ 较多使用缩略词，常见的如 PCM（pulse-coded modulation，脉冲编码调制）, CDMA（code division multiple access，码分多址）, DSP（digital signal processing，数字信号处理）等。

④ 词性变换多，例如，sound 一词做名词时，常译为“声音、语音”，做动词时，常译为“听起来”，做形容词时，以“合理的，健全的”较为多见。

2．词法

① 常用一般现在时态，表示真理或客观规律的陈述。

[例 1] Vector and matrix techniques *provide* the framework for much of the developments in modern engineering.

矢量和矩阵方法为现代工程学的发展提供了框架。

② 广泛使用被动语态，强调所论述的客观事物。

[例 2] Chapters 7 and 8 *are devoted to* the calculus of functions of one variable and, recognizing again the mixed background knowledge in mathematics of the students, the basic ideas and techniques of differentiation and integration *are reviewed* in Chapter 7.

第 7 章和第 8 章讨论单变量函数的微积分，考虑到学生数学基础参差不齐，第 7 章复习微分与积分的基本概念与方法。

③ 普遍使用名词词组及名词化结构，强调客观存在的事实而非某一行为，故常使用表示动作或状态的抽象名词。

[例 3] Television is *the transmission and reception of* images of moving objects by radio waves.

电视通过无线电波发射和接收移动物体的图像。

④使用非限定动词，使句子简明。

[例 4] The calculus, *aided* by analytic geometry, proved *to be* astonishingly powerful and capable of *attacking* hosts of problems that had been baffling and quite unassailable in earlier days.

微积分辅以解析几何是一个非常强大的工具，能够解决许多困扰已久甚至以前认为无法解决的问题。

3．句法

① 较常使用“无生命主语+及物动词+宾语（+宾语补足语）”句型。

[例 5] Chapter 6 *provides* a basic introduction to the ideas of sequences, series and limits.

第 6 章介绍序列、级数及极限等基本概念。

② 常用 it 做形式主语或形式宾语。

[例 6] *It* has been proved that induced voltage causes a current to flow in opposition to the force producing it.

已经证明，感应电压使电流的方向与产生电流的磁场力方向相反。

[例 7] The invention of radio has made *it* possible for mankind to communicate with each other over a long distance.

无线电的发明使人类有可能进行远距离通信。

③ 尽量用紧缩型状语从句而不用完整句。

[例 8] While *designed* primarily for use by engineering students, it is believed that the book is also highly suitable for students of the physical sciences and applied mathematics.

尽管本书主要为工科学生所用，我们相信，它也非常适合于修读物理与应用数学的学生。

④ 割裂修饰比较普遍（包括短语或从句被分隔）。

[例 9] It is hoped that this *provision*, together with the large number of worked examples and style of presentation, also *makes* the book suitable for private or directed study. (主谓分离)

希望这些练习，以及书中提供的大量实例及本书的写作风格也使本书适于自学和课堂教学。

⑤ 较多使用祈使语气。

[例 10] *Let* the forward-pass transfer function be given by the linear difference equation.

设前向传递函数由线性差分方程给出。

⑥ 句中并列成分（各种并列短语、单词或从句）较多。

[例 11] Radar has certain inherent advantages over detection systems employing light waves: (1) it has greater range, (2) it is usable in any weather and in day or night, and (3) the electronic circuitry and components for *transmitting, receiving, amplifying, detecting and measuring* are highly developed.

与光波检测系统相比，雷达具有如下优点：（1）检测范围广；（2）可全天候使用；（3）拥有先进的电子元器件与电子线路，可用于信号的发射、接收、放大、检测和测量。

⑦ 复杂长句多。科学技术要阐明事物之间错综复杂的关系，因而需要用复杂的语法关系来表达严密复杂的思维。长句所表达的科技内容严密性、准确性和逻辑性较强。

[例 12] Recognizing the increasing use of numerical methods in engineering practice, which often complement the use of analytical methods in analysis and design and are of ultimate relevance when solving complex engineering problems, there is wide agreement that they should be integrated within the mathematics curriculum.　本句是一个主从复合句，主句为 there is…，由 that 引导同位语从句对 agreement 进行补充说明；现在分词短语 recognizing…构成紧缩型状语从句表原因，which 引导的非限定性定语从句修饰 numerical methods，其谓语为两个并列的成分，该句中还嵌套了一个由 when 引导的紧缩型时间状语从句。

分析与设计过程中常用数值计算方法来弥补解析法的不足，因此在求解复杂的工程问题时，数值方法往往是最为恰当的。由于认识到数值方法在工程实践中的应用日趋增长，人们普遍认为它应该被整合到数学课程中。

Exercises

1. Choose the best answer for each of the following questions.

（1）Mathematical modelling and simulation forms the basis for the analysis and design of engineering systems as a result of ________.

a. developments in computer technology and related software

b. developments in mathematics

c. engineering mathematics

d. the large number of students to study engineering courses

(2) For an engineering student, ________.

a. there is no need to study mathematics

b. he should spend more time on the study of mathematical skills

c. it is enough to learn to use the packaged computer software

d. the main objective is to learn to use mathematics with understanding to solve engineering problem

(3) Which one of the following statements is incorrect according to the text? ________.

a. The exercises are intended to develop skills in mathematical techniques

b. Some of the exercises are open-ended modelling ones which cannot be tackled at the beginning of the course

c. The exercises are only available at the end of each chapter

d. The exercises are designed to develop understanding and encourage learning by doing

(4) Which is not the content of the book? ________

a. differentiation and integration

b. geometry

c. discrete mathematics

d. probability and statistics

(5) ________ is vital to the analysis and design of oscillatory and nonlinear systems.

a. Fourier series analysis

b. Numerical analysis

c. Simulation and modelling

d. Random process

2. Match the term in Column A with the appropriate explanation in Column B.

Column A	**Column B**
(1) mathematical analysis	a. symbolized arithmetic
(2) algebra	b. interconnection between variables
(3) calculus	c. a rectangular array of numeric or algebraic quantities subject to mathematical operations
(4) function	d. a statement of equality between two equal numbers or number systems
(5) probability theory	e. the study of variables and their relationships
(6) matrix	f. the branch of mathematics concerning points, lines, angles, surfaces and solids
(7) equation	g. mathematical treatment concerned with random events
(8) geometry	h. differentiation and integration

3. Translate the following sentences into Chinese.

（1）The calculus, aided by analytic geometry, proved to be astonishingly powerful and capable of attacking hosts of problems that had been baffling and quite unassailable in earlier days.

（2）Of the many remarkable mathematical discoveries made in the 17th century, unquestionably the most outstanding was the invention of calculus.

（3）A great forward stride was made in 1821, when the French mathematician Augustin Louis Cauchy developed an acceptable theory of limits, and then defined continuity, differentiability, and the definite integral in terms of the limit concept.

（4）Mathematical analysis is one of the most important divisions of higher mathematics; its main object is studying variables and their relationships.

（5）The main purpose of a natural or technical science is to establish the relationships between the variables involved in the process under consideration and to describe it mathematically.

（6）Mathematical methods lie in the foundation of physics, mechanics, engineering and other natural sciences. For all of them mathematics is a powerful theoretical and practical tool without which no scientific calculation and no engineering and technology are possible.

Reading Material

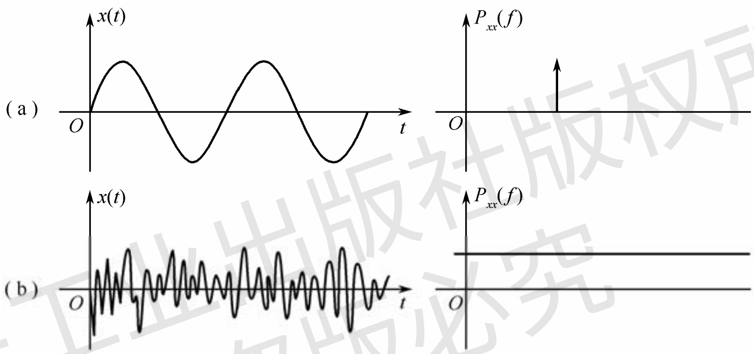
Fourier Analysis and Synthesis

Jean Baptiste Joseph Fourier (1768—1830) studied the mathematical theory of heat conduction in his major work, *The Analytic Theory of Heat*. He established the partial differential equation governing heat diffusion and solved it using an infinite series of trigonometric functions. The description of a signal in terms of elementary trigonometric functions had a profound effect on the way signals are analyzed. The Fourier method is the most extensively applied signal-processing tool. This is because the transform output leads itself to easy interpretation and manipulation, and leads to the concept of frequency analysis. Furthermore even biological systems such as the human auditory system perform some form of frequency analysis of the input signals. The applications of the Fourier transform include filtering, telecommunication, music processing, pitch modification, signal coding and signal synthesis, feature extraction for pattern identification as in speech recognition, image processing, spectral analysis in astrophysics and radar signal processing.

1. Introduction

The objective of signal transformation is to express a signal as a combination of a set of basic “building block” signals, known as the basis functions. The transform output should lead itself to convenient analysis, interpretation and manipulation. A useful consequence of transforms, such as the Fourier and the Laplace, is that differential analysis on the time domain signal becomes simple algebraic operations on the transformed signal. In the Fourier transform the basic building block signals are sinusoidal signals with different periods giving rise to the concept of frequency. In Fourier analysis a signal is decomposed into its constituent sinusoids, i.e., frequencies, the amplitudes of various frequencies form the so-called frequency spectrum of the signal. In an inverse Fourier transform operation the signal can be synthesized by adding up its constituent frequencies. It turns out that many signals that we encounter in daily life such as speech, car engine noise, bird songs, music etc. have a periodic or quasi-periodic structure, and that the cochlea in the human hearing system performs a kind of harmonic analysis of the input audio signals. Therefore the concept of frequency is not a purely mathematical abstraction in that biological and physical systems have also evolved to make use of the frequency analysis concept.

The power of the Fourier transform in signal analysis and pattern recognition is its ability to reveal spectral structures that may be used to characterize a signal. This is illustrated in Figure 1 for the two extreme cases of a sine wave and a purely random signal. For a periodic signal the power is concentrated in extremely narrow bands of frequencies indicating the existence of structure and the predictable character of the signal. In the case of a pure sine wave as shown in Figure 1(a) the signal power is concentrated in one frequency. For a purely random signal as shown in Figure 1(b) the signal power is spread equally in the frequency domain indicating the lack of structure in the signal.



(a) a predictable signal (b) a random signal

Figure 1 The concentration or spread of power in frequency indicates the correlated or random character of a signal

2. Fourier Series: Representation of Periodic Signals

The following three sinusoidal functions form the basis functions for the Fourier analysis

$$x_1(t) = \cos\omega_0 t \tag{1}$$

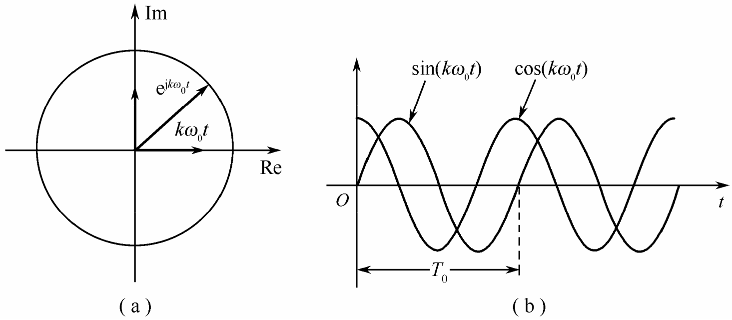
$$x_2(t) = \sin\omega_0 t \tag{2}$$

$$x_3(t) = \cos\omega_0 t + \mathrm{j}\sin\omega_0 t = \mathrm{e}^{\mathrm{j}\omega_0 t} \tag{3}$$

Figure 2(a) shows a vector representation of the complex exponential in a complex plane with real (Re) and imaginary (Im) dimensions, and Figure 2(b) shows the cosine and the sine components of the complex exponential (cissoidal) signal of Eq.(3). The Fourier basis functions are periodic with an angular frequency of ω_0 rad/s and a period of $T_0 = 2\pi/\omega_0 = 1/F_0$ seconds, where F_0 is the frequency in Hz. The following properties make the sinusoids the ideal choice as the elementary building block basis functions for signal analysis and synthesis:

（i）Orthogonality; two sinusoidal functions of *different* frequencies have the following orthogonal property:

$$\int_{-\infty}^{\infty} \sin\omega_1 t \sin\omega_2 t \mathrm{d}t = -\frac{1}{2}\int_{-\infty}^{\infty} \cos(\omega_1+\omega_2)t\mathrm{d}t + \frac{1}{2}\int_{-\infty}^{\infty} \cos(\omega_1-\omega_2)t\mathrm{d}t = 0 \tag{4}$$



(a) vector representation of a complex exponential; (b) real and imaginary parts of a complex sinusoid

Figure 2 Fourier basis functions

For harmonically related sinusoids the integration can be taken over one period. Similar equations can be derived for the product of cosines, or sine and cosine, of different frequencies. Orthogonality implies that the sinusoidal basis functions are independent and can be processed independently. For example in a graphic equalizer we can change the relative amplitudes of one set of frequencies, such as the bass, without affecting other frequencies, and in subband coding different frequency bands are coded independently and allocated different number of bits.

（ii）Sinusoidal functions are infinitely differentiable. This is important, as most signal analysis, synthesis and manipulation methods require the signals to be differentiable.

（iii）Sine and cosine signals of the same frequency have only a phase difference of π/2 or equivalently a relative time delay of a quarter of one period, i.e., $T_0/4$.

Associated with the complex exponential function $\mathrm{e}^{\mathrm{j}\omega_0 t}$ is a set of harmonically related complex exponential of the form

$$\left[1, \mathrm{e}^{\pm\mathrm{j}\omega_0 t}, \mathrm{e}^{\pm\mathrm{j}2\omega_0 t}, \mathrm{e}^{\pm\mathrm{j}3\omega_0 t}, \cdots\right] \tag{5}$$

The set of exponential signals in Eq.(5) are periodic with a fundamental frequency $\omega_0 = 2\pi/T_0 = 2\pi F_0$ where T_0 is the period and F_0 is the fundamental frequency. These signals form the set of basis functions for the Fourier analysis. Any linear combination of these signals of the form

$$\sum_{k=-\infty}^{\infty} c_k \mathrm{e}^{\mathrm{j}k\omega_0 t} \tag{6}$$

is also periodic with a period of T_0. Conversely any periodic signal $x(t)$ can be synthesized from a linear combination of harmonically related exponentials.

The Fourier series representation of a periodic signal is given by the following synthesis and analysis equations:

$$x(t) = \sum_{k=-\infty}^{\infty} c_k \mathrm{e}^{\mathrm{j}k\omega_0 t} \qquad k=\cdots-1,0,1,\cdots \qquad \text{Synthesis equation} \tag{7}$$

$$c_k = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\mathrm{e}^{-\mathrm{j}k\omega_0 t}\mathrm{d}t \quad k=\cdots-1,0,1,\cdots \qquad \text{Analysis equation} \tag{8}$$

The complex-valued coefficient c_k conveys the amplitude (a measure of the strength) and the phase of the frequency content of the signal at $k\omega_0$ Hz. Note from the analysis Eq.(8), that the coefficient c_k may be interpreted as a measure of the correlation of the signal $x(t)$ and the complex exponential $\mathrm{e}^{-\mathrm{j}\omega_0 t}$.

The set of complex coefficients $\cdots c_{-1}, c_0, c_1 \cdots$ are known as the signal spectrum. Eq.(7) is referred to as the synthesis equation, and can be used as a frequency synthesizer (as in music synthesizers) to generate a signal as a weighted combination of its elementary frequencies. The representation of a signal in the form of Eq.(7) as the sum of its constituent harmonics is also referred to as the *complex Fourier* series representation. Note from Eq.(7) and (8) that the complex exponentials that form a periodic signal occur only at discrete frequencies which are integer multiples, i.e., harmonics, of the fundamental frequency ω_0. Therefore the spectrum of a periodic signal, with a period of T_0, is *discrete* in frequency with *discrete spectral lines* spaced at integer multiples of $\omega_0 = 2\pi / T_0$.

3. Fourier Transform: Representation of Aperiodic Signals

The Fourier series representation of periodic signals consists of harmonically related spectral lines spaced at the integer multiples of the fundamental frequency. The Fourier representation of aperiodic signals can be developed by regarding an aperiodic signal as a special case of a periodic signal with an infinite period. If the period of a signal is infinite, then the signal does not repeat itself and is aperiodic. Now consider the discrete spectra of a periodic signal with a period of T_0, as shown in Figure 3(a). As the period T_0 is increased, the fundamental frequency $F_0 = 1/T_0$ decreases, and successive spectral lines become more closely spaced. In the limit as the period tends to infinity (i.e., as the signal becomes aperiodic) the discrete spectral lines merge and form a continuous spectrum. Therefore the Fourier equations for an aperiodic signal, (known as the Fourier transform), must reflect the fact that the frequency spectrum of an aperiodic signal is continuous. Hence to obtain the Fourier transform relation the discrete-frequency variables and operations in the Fourier series should be replaced by their continuous-frequency counterparts. That is the discrete summation sign $\sum$ should be replaced by the continuous summation integral $\int$, the discrete harmonics of the fundamental frequency kF_0 should be replaced by the continuous frequency variable *f*, and the discrete frequency spectrum c_k must be replaced by a continuous frequency spectrum say $X(0)$. The Fourier synthesis and analysis equations for aperiodic signals, the so-called *Fourier transform pair*, are given by

$$x(t) = \int_{-\infty}^{\infty} X(f)\mathrm{e}^{\mathrm{j}2\pi ft}\mathrm{d}f \qquad (9)$$

$$X(f) = \int_{-\infty}^{\infty} x(t)\mathrm{e}^{-\mathrm{j}2\pi ft}\mathrm{d}t \qquad (10)$$

Note from Eq.(10), that $X(f)$ may be interpreted as a measure of the correlation of the signal $x(t)$ and the complex sinusoid $\mathrm{e}^{-\mathrm{j}2\pi ft}$.

The condition for existence and computability of the Fourier transform integral of a signal $x(t)$ is that the signal must have finite energy.

$$\int_{-\infty}^{\infty} |x(t)|^2 \, \mathrm{d}t < \infty \tag{11}$$

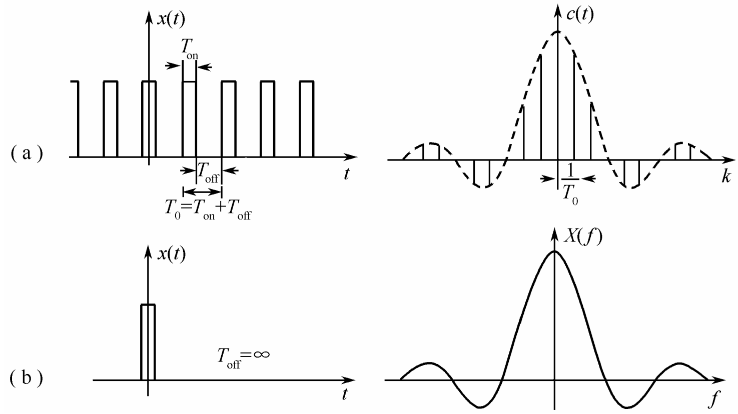


Figure 3 (a) A periodic pulse train and its line spectrum, (b) a single pulse from the periodic train in (a) with an imagined 'off' duration of infinity; its spectrum is the envelope of the spectrum of the periodic signal in (a).

New Words

aperiodic	非周期的
auditory	耳的，听觉的
cochlea	耳蜗
coefficient	系数
computability	可计算性
constituent	作为整体一部分的；组成的
correlation	相互关系，相关（性）
cosine	余弦
equalizer	均衡器
filtering	滤波（filter，滤波器）
harmonic	谐波，谐函数
orthogonality	正交性，正交状态
product	乘积
random	随机的
sinusoidal	正弦曲线
synthesis	综合，合成
quasi-periodic	拟周期的

Questions

1．Give a simple example of the use of Fourier analysis.

2．Why do we choose sinusoids as basic building block functions for signal analysis and synthesis?

3．What is the difference between the Fourier series and the Fourier transform?

Unit 2

Text

Simulation of Random Variables

Many events in our lives are subject to chance — by which we mean that they are not entirely predictable. To some extent, we can choose where we live and what sort of work we do, but even so we cannot be sure what sort of neighbors or workmates we shall have: noisy, generous, friendly and so on. In a similar way, experiments in all branches of science and engineering involve unpredictable outcomes that may be expressed either as a quality such as ‘turned green’ or ‘exploded’, or numerically in terms of mass, resistance or any standard unit. In contrast with everyday life, an ‘experiment’ is repeated many times, so that the limited predictability of the various outcomes can emerge as a pattern within the disorder. The subject of statistics is about extracting that pattern and drawing useful conclusions from it, and the theoretical foundation for this is contained in the theory of probability.

2.1 Interpretations of Probability

The theory of probability underlies the methods of inference used in statistical situations, and the concept of probability can be related to the histogram of data. The height of each bar determines the proportion of the sample that fell into the corresponding class. One way to think of probability is to assume that as a larger and larger sample is taken, the histogram will stabilize and the class proportions will converge to the ‘true’ probability figures. *This concept of probability is of an objective quantity that applies to each observation and measures (in a relative way) how likely it is to fall into the corresponding class*[1]. Like the speed of light and the density of gold, it is known only imperfectly because of our limited capacity to do experiments.

An alternative concept of probability that is important in decision-making and expert systems involves **degree of belief**. This is highly subjective, because it will depend upon the individual (or group) concerned and will vary with past experience. This seems unscientific at first sight, and there is much resistance to this notion, but there are many situations where experiments are unrepeatable in principle and no ‘large-sample proportion’ approach is applicable. The outcome of an election is uncertain, and it is not unreasonable to say that some outcomes are ‘more probable’ than others, but the actual election can take place only once. *It seems that one is forced into a subjective view of the uncertainties, but the probability figures that emerge must obey certain rules in order to be consistent*[2]. Advocates of subjective probability have shown that these rules are the same as those obeyed by the sample proportions.

The formal theory of probability admits a number of ‘interpretations’, of which these objective and subjective interpretations are by far the most important. *For engineering students it is most appropriate to keep the first interpretation — that of probability as an idealized proportion — in mind when studying the theory*[3].

2.2 Random Variables

2.2.1 Definition

A **random variable** consists of a sample space of possible numerical values together with a probability distribution over those values.

Random variables vary in their degree of advance predictability. As the following four examples show, the probabilities of the possible values are very dispersed for some random variables, but highly concentrated for others:

(a) *The toss of a die.* No die is perfect, but for this random variable the probability of the six values are almost equal.

(b) *Next month's rainfall.* Unless you live in a part of the world that has a very constant climate, the amount of rain that falls in March, say, varies from year to year quite considerably. The probabilities are not quite so dispersed as for the die toss, but there is a high degree of uncertainty.

(c) *A flight delay.* Here there is a high probability of at most a short delay, but a small probability of a very long delay. The probabilities are relatively concentrated.

(d) *The time of tomorrow's sunrise.* Knowing your latitude, longitude, altitude, the date, the direction of sunrise and the height above sea level of the horizon in that direction, you could predict the time very precisely. There would be some small uncertainty because of atmospheric refraction.

The behavior of a random variable is determined by the profile of its probability distribution. We shall now enlarge upon this for the two common types. The notation convention is to denote a random variable by a capital letter, say X, and an observed value by the corresponding lower-case letter, then x.

2.2.2 Discrete Random Variables

The random variable X, say, has a list of possible values $v_1, v_2, \cdots, v_m$ with probabilities $P(X = v_1), \cdots, P(X = v_m)$ of equaling these values. In other words, each actual value x of X is equal to v_i for some $i = 1, \cdots, m$, and we allow m to be infinite if required. Typical examples are die tosses, birthdays, and the numbers of defective components in a batch from a production line.

In general, the behavior of a discrete random variable can be represented graphically by means of a **probability function**

$$P_X(x) = P(X = x) \quad (-\infty < x < +\infty)$$

Also useful is the **distribution function** $F_X(x)$ defined as

$$F_X(x) = P(X \leqslant x) \quad (-\infty < x < +\infty)$$

2.2.3 Continuous Random Variables

A continuous random variable X can take any value within some interval (v_1, v_2). If this interval is not already infinite, we define the random variable to be zero outside it and extend the domain of definition to $(-\infty, +\infty)$. Typical examples are a person's height and weight, component lifetimes, and all measured quantities expressed in units of mass, length, time, temperature, resistance and so on.

In general, the behavior of a continuous random variable X is described by a probability density function $f_X(x)$ for $(-\infty < x < +\infty)$ as illustrated in Figure 2.1. As will be explained below, $f_X(x)$ is not the probability that $X = x$; instead, the density function has to be understood in terms of the **distribution function** $F_X(x)$, which measures (as before) the probability that the value of the random variable is less than or equal to the argument x:

$$F_X(x) = P(X \leqslant x) \quad (-\infty < x < \infty)$$

In this case, because there are no discrete steps in probability, $F_X(x)$ is continuous and differentiable, and its derivative is called the probability density function $f_X(x)$:

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}[F_X(x)]$$

The significance of the density function is that it indicates for a continuous random variable the concentration of possible observed values along the real axis.

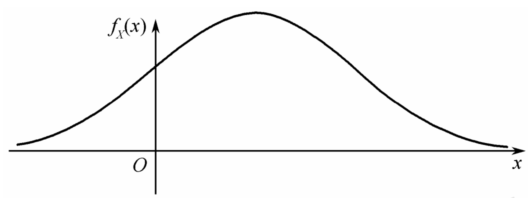


Figure 2.1 A typical probability density function

2.3 Random Variables for Simulation

Computer simulations are very widely used in research, design and training. Perhaps the best known is the flight simulator upon which pilots receive much of their training. Simulations are used in research and design wherever a system is too complex for a complete solution to a problem to be obtained theoretically, or where a solution can be obtained but its completeness or accuracy is open to question.

Simulations are deterministic if what occurs at any time is completely determined by the state of the system. In contrast, they are stochastic if what occurs at any time can be influenced by a chance element that is inherently unpredictable. Stochastic simulations therefore require that random variables (or outcomes) be generated within the program. *This may seem a hopeless requirement, considering that computer programs are sequences of deterministic instructions running on deterministic hardware*[4]. However, it is possible to generate sequences of numbers that are deterministic and repeatable but that have the appearance of being random. These pseudo-random numbers are very useful for simulations, and for other purposes such as the so-called Monte Carlo numerical methods.

Most modern computers contain a software facility for generating pseudo-random numbers with a uniform distribution on the interval (0, 1):

$$f_U(u) = \begin{cases} 1 & (0 < u < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

The successive variables $\{U_1, U_2, \cdots\}$ *appear to be uncorrelated, and, although there is some structure in the sequence (and indeed the sequence will eventually repeat itself), it is rare for these* *deficiencies to cause problems in practice*[5].

Random variables with non-uniform distributions are obtained from the sequence $\{ U_1, U_2, \cdots \}$ by applying various transformations. Figure 2.2 contains pseudo-code listings for generating the most common random variables. In each case it is assumed that the system function “rnd” returns a uniform (0, 1) value, which is stored in the variable U. The variable X contains the required value of the random variable. The binomial is based on the Bernoulli, the Poisson on the exponential, and the normal on the central limit theorem.

{**Bernoulli random variable** X, parameter p.}
$U \leftarrow$ rnd
if $U < p$ then $X \leftarrow 1$ else $X \leftarrow 0$ endif
{**Binomial random variable** X, parameters n, p.}
$X \leftarrow 0$
for i is 1 to n do
    $U \leftarrow$ rnd
    if $U < p$ then $X \leftarrow X + 1$ endif
endfor
{**Exponential random variable** X, parameter L, uses log function to base e.}
$U \leftarrow$ rnd; $X \leftarrow -(\log(u))/L$
{**Poisson random variable** X, parameter L.}
$X \leftarrow -1; W \leftarrow 1; P_0 \leftarrow \exp(-L)$
repeat
    $X \leftarrow X + 1$; $U \leftarrow$ rnd; $W \leftarrow W * U$
until $W < P_0$
{**Normal random variable** X, parameters mean, sd.}
$T \leftarrow 0$
for i is 1 to 12 do
    $U \leftarrow$ rnd; $T \leftarrow T + U$
endfor
$X \leftarrow$ sd $* (T - 6) +$ mean

Figure 2.2 Pseudo-code listings for non-uniform random variables

Words & Expressions

advocate	[ˈædvəkit]	*n.*	主张者，赞成者
alternative	[ɔːlˈtəːnətiv]	*adj.*	另一个可选择的
argument	[ˈɑːgjumənt]	*n.*	自变量
concentration	[ˌkɔnsenˈtreiʃən]	*n.*	集中，集合，浓缩
converge	[kənˈvəːdʒ]	*v.*	收敛（*n.* convergence）
deficiency	[diˈfiʃənsi]	*n.*	缺乏，不足
derivative	[diˈrivətiv]	*n.*	导数
die	[dai]	*n.*	骰子
differentiable	[ˌdifəˈrenʃiəbl]	*adj.*	可微分的
disperse	[disˈpəːs]	*v.*	（使）分散，（使）散开，疏散
facility	[fəˈsiliti]	*n.*	设备，工具
histogram	[ˈhistəugræm]	*n.*	直方图
inference	[ˈinfərəns]	*n.*	推论

interpretation	[in'tə:prəteiʃən]	*n.*	说明，解释
profile	['prəufail]	*n.*	剖面，侧面，外形，轮廓
proportion	[prə'pɔ:ʃən]	*n.*	比例
pseudo-random	['psju:dəu 'rændəm]	*adj.*	伪随机的
refraction	[ri'frækʃən]	*n.*	折射
stabilize	['steɪbɪlaɪz]	*v.*	稳定（*n.* stabilization）
statistical	[stə'tistikəl]	*adj.*	统计的，统计学的
stochastic	[stəu'kæstik]	*adj.*	随机的
successive	[sək'sesiv]	*adj.*	连续的
toss	[tɔs]	*v.*; *n.*	投，掷
underlie	[ˌʌndə'lai]	*v.*	位于……之下，成为……的基础
by means of			依靠
consist of			由……组成
give rise to			引起，使发生
in contrast with			和……形成对比［对照］
in terms of			根据，按照，用……的话，在……方面
Bernoulli distribution			伯努利分布
Binomial distribution			二项式分布
Normal distribution			正态分布
Poisson distribution			泊松分布
probability density function			概率密度函数（*abbr.* pdf）
Uniform distribution			均匀分布

Notes

[1] This concept of probability is of an objective quantity that applies to each observation and measures (in a relative way) how likely it is to fall into the corresponding class. “be of …”这种结构用来表示人或物的特征，通常译为“具有……”。

概率的这种定义是客观的，它应用于每次观测，并以一种相对的方式度量其结果并归入相应类别的可能性。

[2] It seems that one is forced into a subjective view of the uncertainties, but the probability figures that emerge must obey certain rules in order to be consistent. “that emerge”是定语从句，修饰先行词 figures（数字）。

看来，人们对这种不确定性陷入了自己的主观判断中，但是由此产生的概率必须遵循某些准则以保持一致性。

[3] For engineering students it is most appropriate to keep the first interpretation — that of probability as an idealized proportion — in mind when studying the theory. “that of …”补充说明 interpretation，that 即指 interpretation，它们将 keep … in mind 分隔。

对于工科学生，学习概率论时应谨记，概率就是理想的比例。

[4] This may seem a hopeless requirement, considering that computer programs are sequences of

deterministic instructions running on deterministic hardware. “considering that …”引导原因状语从句。

由于计算机程序是在确定的硬件上运行的确定的指令，这似乎是个奢求。

[5] The successive variables $\{U_1, U_2, \cdots\}$ appear to be uncorrelated, and, although there is some structure in the sequence (and indeed the sequence will eventually repeat itself), it is rare for these deficiencies to cause problems in practice. “the successive variables”是指序列中的变量，强调个体，而“the sequence”则是指$\{U_1, U_2, \cdots\}$这个序列，强调整体。

序列$\{U_1, U_2, \cdots\}$中的元素看似互不相关，尽管在序列中存在某些结构（该序列最终的确会重复），但这并不影响实际应用。

Grammar

专业英语词汇的构成

专业英语与普通英语在词汇方面表现出很大的差异，许多单词在普通英语中的意义与在专业英语中的意义大相径庭，同一个词在不同专业的意义也不尽相同。一般来讲，专业英语词汇由三部分组成：专业词汇、半专业词汇和非专业词汇（或普通词）。专业词汇（technical words）主要指在某一学科、某一专业所独有的专用术语，只有一种专业含义，非常单纯，如 histogram（直方图），probability（概率），oscillator（振荡器）等 。半专业词汇（semi-technical words）是指跨学科出现的频率很高的词，在不同的专业领域具有不同的精确含义，如 carrier 在电学中表示“载波，载流子”，在机械领域表示“载重架”，在化学中表示“载体，吸收剂”，在医学中表示“带菌者”等。非专业词汇（non-technical words）是指在非专业英语中很少使用，却严格属于非专业英语性质的词汇，如 application, implementation, to yield 等，这些词也包括出现频率高、在语法上起重要作用的结构功能词，如限定词、介词、连词等。

专业英语词汇中有很大一部分符合构词法。常用构词法主要包括：派生、复合、转化、拼缀和缩略等，下面分别予以介绍。

1. 派生法（Derivation）

派生法通过在原有词或词根的基础上加前缀或（和）后缀而构成新词，前缀通常用以修饰或改变词意，后缀显示词性。表 2.1～表 2.4 分别列出电子与通信专业常用的前、后缀及词根。

表 2.1 常用前缀

词 缀	意 义	词 例
a-	不，无	asymmetry, asynchronous
anti-	反抗，防止	anti-clockwise
co-	共同，相互	cooperation, correlation
dis-	不，除	disadvantage, discover
en-, em-	使	enable, enlarge, embed, embody
im-, il-, in-, -ir	不	imbalance, illegible, incorrect, irregular
inter-	在……之间，相互	interchange, interface, internet, interact
mis-	不，失	miscount, mistake

词　　缀	意　　义	词　　例
multi-, poly-	多	multipurpose, polynomial
non-	非，不	nonlinear, nondestructive
post-	在后	postgraduate, posterior
pre-	预先	preset, preface
re-	再，反，重新	reaction, readjust, reverse
sub-	在下，次于	subroutine, subscript
super-	在上，超	superconductor, superposition
sym-, syn-	相同	symmetry, synchronous
tele-	远离	telephone, telegraph
trans-	跨，移	transmitter, transverse, transform
ultra-	外，极，超	ultrasonic, ultraviolet ray
un-	不	unable, unavoidable, unstable

表 2.2　表示数量关系的常用前缀

词缀	意义	符号	词例	词缀	意义	词例
pico-	10^{-12}	p	picofarad	semi-, hemi-	半	Semiconductor, hemisphere
nano-	10^{-9}	n	nanometer	uni-, mon-	一	monotone, uniform
micro-	10^{-6}	μ	microhenry	bi-, di-, ambi-, twi-	二	bilateral, diode, ambiguous, twilight
milli-	10^{-3}	m	millisecond	tri-	三	triangle, tripod
kilo-	10^{3}	k	kilogram	quadr-	四	quadruple
mega-	10^{6}	M	megahertz	oct-	八	octagon
giga-	10^{9}	G	gigabytes	deca-	十	decade

表 2.3　常用后缀

作用	词　　缀	意　　义	词　　例
名词词尾	-ance, -ancy, -cy, -ence, -ency	表示情况、性质、状态、程度等	inference, resistance, efficiency, accuracy
	-er,or	表示人或物	amplifier, conductor, researcher
	-ic(s)	学（科），学术	logic, electronics
	-tion, -sion	表示动作及其过程、状态和结果	distribution, conclusion
	-ing		readings, recordings
	-ist	表示人	scientist, specialist
	-(i)ty	表示性质、程度等	probability, uncertainty
	-ment	表示动作、状态等	measurement, development
	-ture	表示性质、状态等	mixture, temperature
	-ness	形容词→名词后缀	hardness, robustness
形容词词尾	-able	表示可能性	differentiable, countable, controllable
	-(c)al, -ic(al)	表示性质，……的	statistical, atomic, typical
	-ant, -ent	表示状态、性质等	convergent, important, independent
	-ar(y)	与……有关的	circular, secondary
	-ed	已……的，被……的	reduced, treated, refined, developed
	-ive	表示性质、状态等	objective, relative, effective
	-ful	充满……的	plentiful, useful
	-less	没有，无……的	useless, countless
	-ous	有……性质的	numerous, various

续表

作用	词　缀	意　义	词　例
动词词尾	-en	使变成……	harden, broaden
	-ize, -ise	使成为，……化	modernize, stabilize
	-fy	使成为，……化	classify, verify
副词词尾	-ly	地，每……（一次）地	closely, likely, imperfectly, monthly
	-ward(s)	表示方向	backwards, upward(s)
	-wise	表示方式、方向	clockwise, likewise

表 2.4　常用词根

词　根	意　义	词　例
audi-	听，听见	audibility, audiphone
auto-	自动，自己	automation, autopilot
-free	无	rustfree
-fold	倍	multifold
-gram, -graph	记录物，写，文字，图形等	spectrogram, telegram, spectrograph
-graphy	图像学	photography
-meter	仪表，仪器	tachometer
-ology	学（科）	biology, geology
-phone	声音	microphone, telephone
-proof	防……	waterproof
-scope	观测仪	telescope
-tight	密，不透	airtight

2. 复合法（Composition）

复合法是指由两个或两个以上的词按照一定的次序排列，以构成新词。多数复合词可通过其组成部分猜测词意。例如，trial and error（反复试验），Q-factor（品质因子），allowable error（允许误差）等。此外，常采用以下两种形式构成复合词：

- 直接结合，如 breakthrough（突破），overestimate（高估），bandwidth（带宽）等；
- 用连字符结合，如 general-purpose（多种用途的），state-of-the-art（达到最新技术发展水平的）等。

3. 转化法（Conversion）

转换法即单词词性转换，词性转换后其意义与原意有着密切联系，如 function, sound, ground 等词。有些词还会发生音变，如 use, record 等；或音移，如 increase, research, subject 等。

4. 拼缀法 (Blending)

拼缀法以原有的两个或两个以上的词为基础，经过首尾剪裁（或保留其中一个原词），重新组合而成，是复合词的缩略形式，如 transistor（晶体管）=*tran*sfer + re*sistor*, modem（调制解调器）由 *mo*dulator 和 *dem*odulator 拼缀而成。

5. 缩略法（Shorting）

缩略法是将几个单词的首字母以大写形式缩合到一起成为一个新词。该法多用于专有名词，利于记忆，如 radar（RAdio Detection And Ranging，雷达），GPS（Global Positioning System，全球定位系统）等。

随着电子信息及通信类专业技术日新月异的发展，不断有新的专业词汇出现。大多数新词都是用以上五种传统的构词法构成，掌握了基本词汇，运用构词知识，一方面可提高记忆效率，扩大词汇量，另一方面还可以逆向利用构词知识分解单词，了解单词的来历及新添含义，消除阅读障碍。

Exercises

1. Choose the best answer for each of the following questions.

（1）The theoretical foundation for the subject of statistics is contained in ________.

a. histogram

b. the concept of probability

c. the theory of probability

d. decision-making and expert system

（2）For engineering students it is most appropriate to keep the first interpretation — *that* of probability as an idealized proportion — in mind when studying the theory. What does “that” mean in the sentence? ________.

a. theory

b. concept

c. interpretation

d. study

（3）Find which is not an example of random variables ________.

a. the number of students attending class

b. the number of bits used to denote ASCII codes

c. the lifetime of a battery

d. the waiting time at the crossroad

（4）In general, the behavior of a continuous random variable is described by a ________.

a. distribution function

b. probability function

c. probability distribution

d. probability density function

（5）Random variables with different distributions can be generated ________.

a. by Monte Carlo numerical methods

b. from an Gaussian distributed random sequence

c. from uniformly distributed pseudo-random sequence

d. based on the central limit theorem

2. Find the word pairs, one from column A and one from column B.

Column A	Column B
（1）random	a. space
（2）expert	b. frequency
（3）probability	c. simulation
（4）density	d. deviation
（5）computer	e. function
（6）sample	f. variable
（7）statistical	g. distribution
（8）impossible	h. event
（9）standard	i. system
（10）relative	j. inference

3. Translate the following sentences into Chinese.

（1）A lot of information is required to specify the exact distribution of a random variable, and even more to specify the joint distribution of two or more variables.

（2）The binomial and Poisson are discrete distributions, which have the widest applications among all discrete random variables. The probability distribution is especially useful to engineers because of its importance in statistical quality control.

（3）For any random variable the difference between the values of the distribution function at two points is the probability that a value of the random variable will lie between those two points (or is equal to the upper one).

（4）In general, the probability p of a random event can be interpreted as meaning that if the experiment is repeated a large number of times, the event would be observed about $100p$ percent of the time.

（5）In statistics it is customary to refer to any process of observation as an experiment.

（6）If an event definitely cannot occur upon realization of the set of conditions it is called impossible.

4. Complete the sentences. You may have to change some words slightly.

（1）_____ is the study and use of electrical devices that operate by controlling the flow of _____ or other _____ charged particle in devices such as vacuum tube and semiconductors.

a. electron　　b. electronic　　c. electronics　　d. electronically

（2）One of the best known computer _____ which are very widely used in research, design and training is the flight _____ upon which pilots receive much of their training.

a. simulation　　b. simulator　　c. simulating　　d. simulate

（3）Signals are sent to _____ to _____.

a. amplification　　b. amplifier　　c. amplifying　　d. amplify

（4）

（a）It follows that we can use the one table for the standard normal for all calculations involving normal _____.

(b) However, a great many applications of probability theory concern quantitative _____ rather than qualitative events.

(c) There are two approaches to measuring the _____ of random variables around their central values. The most important such measure is the _____, a weighted sum of sequential differences between the possible values and the mean.

(d) Frequency response will be flat and bandwidth infinite because AC will be simply a rapidly _____ DC level to the ideal amplifier.

a. various b. variable c. variate d. variance e. variation f. varying

Reading Material

The Central Limit Theorem

1. The Normal Distribution

One class of distributions is awarded the name 'normal' because of the regularity with which random continuous data are found to obey it. This is no coincidence. The central limit theorem provides an explanation in terms of cumulative independent random parts adding up to a normal whole, a situation that will be of great value in statistical inference. The normal distribution also serves as an approximation to the binomial distribution that complements the Poisson approximation.

The normal distribution has two parameters, which can be shown to be the mean and standard deviation, so the appropriate symbols μ_x and σ_x are used.

A continuous random variable X has a **normal distribution** with mean μ_x and variance σ_x^2 if

$$f_X(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right] \qquad (-\infty < x < +\infty, \sigma_x > 0)$$

This density function is symmetrical about μ_x and has the bell-shaped form shown in Figure 1. This distribution is also sometimes referred to by its more traditional name: the **Gaussian distribution.**

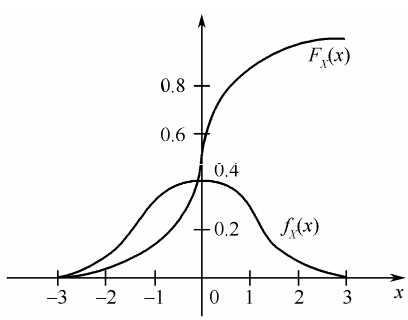


Figure 1 The normal density and distribution (for $\mu_x = 0$ and $\sigma_x = 1$)

The need to declare that a random variable has a normal distribution (with a specified mean and variance) is so common that a special notation exists for the purpose:

$$X \sim N(\mu_x, \sigma_x^2)$$

Calculations involving the normal distribution are complicated by the fact that there is no simple expression for the integral of the density function on an arbitrary interval; in other words, the distribution function $F_X(x)$ does not have a simple explicit form. Instead, tables of this function are used. In fact, only a single table is needed: that for the special case of a normal distribution with a mean of zero and a variance of one.

The **standard normal** cumulative distribution function is

$$\phi(z) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{z} \mathrm{e}^{-x^2/2} \mathrm{d}x$$

This function is usually tabulated only for $z \geqslant 0$; for $z < 0$ the symmetry implies that

$$\phi(-z) = 1 - \phi(z)$$

A typical table of the standard normal function $\phi(z)$ is usually available. The variance is not changed by this subtraction, but then dividing by the standard deviation gives a variable with a variance of one:

$$\mathrm{Var}\left(\frac{X - \mu_x}{\sigma_x}\right) = 1$$

It is a property of the normal distribution, not shared by most distributions, that the result of this operation is still normal. It is usual to denote the new random variable by the letter Z:

$$Z = \frac{X - \mu_x}{\sigma_x}$$

This is then a standard normal random variable. Conversely, any normal random variable can be considered to have been obtained from a standard normal random variable by multiplying by the required standard deviation and adding the mean:

$$X = \sigma_x Z + \mu_x$$

It follows that we can use the one table for the standard normal for all calculations involving normal variates.

2. The Central Limit Theorem

The practical methods of statistical inference have foundations in probability theory, and the fundamental assumption underlying many of these methods is that the data have a distribution that is normal. Some statistical methods are robust in the sense that they work reliably even under moderate violations of their assumptions, but it is unsatisfactory to rely heavily upon this. If normality of the data were exceptional then this would severely limit the scope of those methods that assume it. Fortunately (and as the name implies), the normal distribution arises very frequently in practice; the reason for this will be explained in this section.

Continuous measurements of random phenomena such as noise in electronic circuits or wave elevation on the sea surface give rise to graphs of the form shown in Figure 2. If the signal is sampled at regular intervals and a histogram of values built up, it is often found that the histogram closely approximates to a normal density curve. Physically, there are many separate independent random components adding up to produce the measured signal, and it is the total that is normal.

There are many sources of noise in an electronic circuit and there are many separate waves on the sea. That the cumulative effect of these, which are often not individually normal, is to produce a total that has that special character is the substance of the following result.

Figure 2 Continuous signal with normal distribution

Theorem 1 (Central limit theorem): If $\{X_1, \cdots, X_n\}$ are independent and identically distributed random variables (the distribution being arbitrary), each with mean μ_x and variance σ_x^2, and if

$$W_n = \frac{X_1 + \cdots + X_n}{n} \quad \text{and} \quad Z_n = \frac{X_1 + \cdots + X_n - n\mu_x}{\sigma_x \sqrt{n}}$$

then as $n \to \infty$, the distributions of W_n and Z_n tend to $W_n \sim N\left(\mu_x, \sigma_x^2 / n\right)$ and $Z_n \sim N(0,1)$ respectively. Or, loosely speaking, the sum of independent identically distributed random variables tends to a normal distribution.

The following points should be noted:

● The standard normal is obtained by subtracting the mean of the total and dividing by the standard deviation.

● The distributions converge to the normal in the sense that the cumulative distribution functions converge. This ensures that all observational properties of Z_n will be standard normal for sufficiently large *n*.

● How large *n* has to be before the normal approximation is good depends upon the underlying population. If the distribution of the variables X_i is symmetric about the mean then convergence to the normal is rapid. Figure 3(a) shows the distributions of the uniform random variable X with density function

$$f_X(x) = \frac{1}{2}\sqrt{\frac{1}{3}} \quad \left(-\sqrt{3} \leqslant x \leqslant \sqrt{3}\right)$$

(which has mean zero and variance one), together with those for Z_2 and Z_4. The normal distribution is also shown. Figure 3(b) shows similar results for the exponential random variable X with density function

$$f_X(x) = \mathrm{e}^{-(x+1)} \quad (x \geqslant -1)$$

(which has mean zero and variance one), together with Z_5 and Z_{25}. Convergence is clearly more rapid for the symmetric distribution.

● The theorem can be generalized so that the random variables X_i do not need to be identically distributed, which is usually not the case in physical situations.

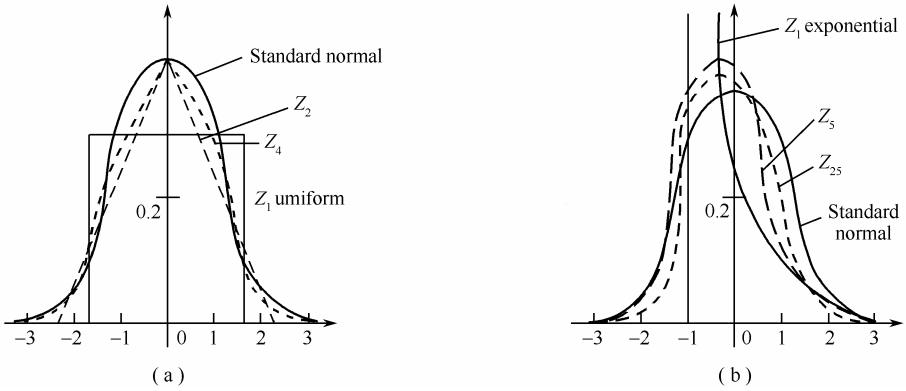


Figure 3 Central limit theorem

(a) uniform; (b) exponential

● Even where the data in an experiment are not normally distributed, the central limit theorem implies that the sample average has a normal distribution for large samples. Much valuable statistics exploits this fact.

New Words

approximation	近似值
arbitrary	任意的
cumulative	累积的
mean	均值
parameter	参数
population	总体
robust	稳定的（-ness，鲁棒性，稳定性）
symmetrical	对称的，均匀的（名词 symmetry）
tabulate	把……制成表格，列表
variance	方差
variate	变量
independent identically distributed	独立同分布（*abbr*. i.i.d.）
standard deviation	标准偏差
standard normal distribution	标准正态分布（均值为 0，方差为 1）
statistical inference	统计推理

Questions

1．Define the standard normal random variable in your own words.

2．Why is it difficult to calculate the cumulative distribution function $F_X(x)$ if x has a normal distribution?

3．What is the use of the central limit theorem in practice?

Unit 3

Text

Circuit Analysis Using the Ideal Operational Amplifier

3.1 Ideal Operational Amplifier

In order to introduce operational amplifier circuitry, we will use an ideal model of the operational amplifier to simplify the mathematics involved in deriving gain expressions, etc., for the circuits presented. With this understanding as a basis, it will be convenient to describe the properties of the real devices themselves in later sections, and finally to investigate circuits utilizing practical operational amplifiers. To begin the presentation of operational amplifier circuitry, then, it is necessary first of all to define the properties of a mythical "perfect" operational amplifier. The model of an ideal operational amplifier is shown in Figure 3.1.

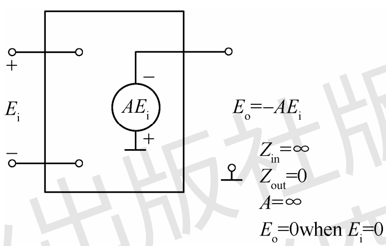


Figure 3.1 Equivalent circuit of the ideal operational amplifier

Defining the Ideal Operational Amplifier

- **Gain:** The primary function of an amplifier is to amplify, so the more gain the better. It can always be reduced with external circuitry, so we assume gain to be infinite.
- **Input Impedance:** Input impedance is assumed to be infinite. This is so the driving source won't be affected by power being drawn by the ideal operational amplifier.
- **Output Impedance:** The output impedance of the ideal operational amplifier is assumed to be zero. It then can supply as much current as necessary to the load being driven.
- **Response Time:** The output must occur at the same time as the inverting input so the response time is assumed to be zero. Phase shift will be 180°. *Frequency response will be flat and bandwidth infinite because AC will be simply a rapidly varying DC level to the ideal amplifier*[1].
- **Offset:** The amplifier output will be zero when a zero signal appears between the inverting and non-inverting inputs.
- **A Summing Point Restraint:** *An important by-product of these properties of the ideal operational amplifier is that the summing point, the inverting input, will conduct no current to the amplifier*[2]. This property is to become an important tool for circuit analysis and design, for it gives us an inherent restraint on our circuit— a place to begin analysis. Later on, it will also be shown that both the inverting and non-inverting inputs must remain at the same voltage, giving us a second

powerful tool for analysis as we progress into the circuits of the next section.

A description of the ideal operational amplifier model was presented in the last section, and the introduction of complete circuits may now begin. *Though the ideal model may seem a bit remote from reality—with infinite gain, bandwidth, etc., it should be realized that the closed loop gain relations which will be derived in this section are directly applicable to real circuits— to within a few tenths of a percent in most cases*[3]. We will show this later with a convincing example.

3.2 The Desirability of Feedback

Consider the open loop amplifier used in the circuit of Figure 3.2. Note that no current flows from the source into the inverting input— the summing point restraint derived in the previous section —hence, there is no voltage drop across R_s and E_s appears across the amplifier input. When E_s is zero, the output is zero. If E_s takes on any non-zero value, the output voltage increases to saturation, and the amplifier acts as a switch.

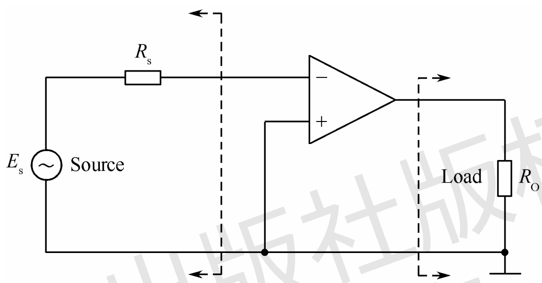


Figure 3.2 Open loop operation

The open loop amplifier is not practical—once an op amp is pushed to saturation, its behavior is unpredictable. Recovery time from saturation is not specified for op amps (except voltage limiting types). It may not recover at all; the output may latch up. The output structure of some op amps, particularly rail-to-rail models, may draw a lot of current as the output stage attempts to drive to one or the other rail.

3.3 Two Important Feedback Circuits

Figure 3.3 shows the connections and the gain equations for two basic feedback circuits. The application of negative feedback around the ideal operational amplifier results in another important summing point restraint: The voltage appearing between the inverting and noninverting inputs approaches zero when the feedback loop is closed.

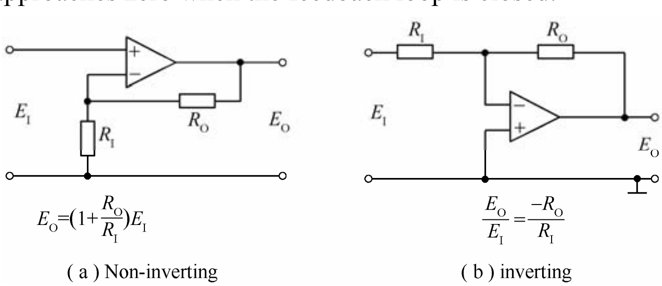


Figure 3.3 Basic amplifier circuits

Consider either of the two circuits shown in Figure 3.3. *If a small voltage, measured at the inverting input with respect to the non-inverting input, is assumed to exist, the amplifier output voltage will be of opposite polarity and can always increase in value (with infinite output available) until the voltage between the inputs becomes infinitesimally small*[4]. When the amplifier output is fed back to the inverting input, the output voltage will always take on the value required to drive the signal between the inputs toward zero.

The two summing point restraints are so important that they are repeated:

● No current flows into either input terminal of the ideal operational amplifier.

● When negative feedback is applied around the ideal operational amplifier, the differential input voltage approaches zero.

These two statements will be used repeatedly in the analysis of the feedback circuits to be presented in the rest of this section.

3.4 Voltage Follower

The circuit in Figure 3.4 demonstrates how the addition of a simple feedback loop to the open loop amplifier converts it from a device of no usefulness to one with many practical applications. Analyzing this circuit, we see that the voltage at the non-inverting input is E_I, the voltage at the inverting input approaches the voltage at the non-inverting input, and the output is at the same voltage as the inverting input. Hence, $E_\mathrm{O} = E_\mathrm{I}$, and our analysis is complete. The simplicity of our analysis is evidence of the power and utility of the summing point restraints we derived and have at our disposal.

Our result also may be verified by mathematical analysis very simply. Since no current flows at the non-inverting input, the input impedance of the voltage follower is infinite. The output impedance is just that of the ideal operational amplifier itself, i.e. zero. Note also that no current flows through the feedback loop, so any arbitrary (but finite) resistance may be placed in the feedback loop without changing the properties of the ideal circuit, shown in Figure 3.5. No voltage would appear across the feedback element and the same mathematical analysis would hold.

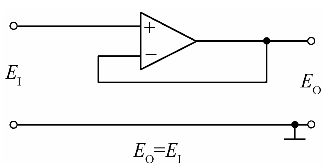


Figure 3.4 Voltage follower

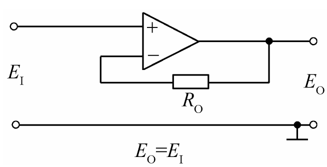


Figure 3.5 Feedback resistor added to the voltage follower

The feedback resistor is of particular importance if the op amp selected is a current-feedback type. The stability of current-feedback op amps is dependent entirely on the value of feedback resistor selected, and the designer should use the value recommended on the data sheet for the device. Unity gain circuits are used as electrical buffers to isolate circuits or devices from one another and prevent undesired interaction. As a voltage following power amplifier, this circuit will allow a source with low current capabilities to drive a heavy load.

The gain of the voltage follower with the feedback loop closed (closed loop gain) is unity.

The gain of the ideal operational amplifier without a feedback loop (open loop gain) is infinity. Thus, we have traded gain for control by adding feedback. Such a severe sacrifice of gain—from infinity to unity—is not necessary in most circuits. The rest of the ideal circuits to be studied will give any (finite) closed loop gain desired while maintaining control through feedback.

3.5 Non-inverting Amplifier

The circuit in Figure 3.3(a) was chosen for analysis next because of its relation to the voltage follower. It is re-drawn as in Figure 3.6, which makes it evident that the voltage follower is simply a special case of the non-inverting amplifier.

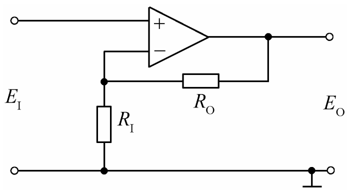


Figure 3.6 Non-inverting amplifier re-drawn to show similarity to the voltage follower

Since no current flows into the inverting input, R_O and R_I form a simple voltage divider. The same voltage must appear at the inverting and non-inverting inputs, so that

$$E_- = E_+ = E_I$$

From the voltage division formula:

$$E_I = \frac{R_I}{R_I + R_O} \cdot E_O$$

$$\frac{E_O}{E_I} = \frac{R_I + R_O}{R_I} = 1 + \frac{R_O}{R_I}$$

The input impedance of the non-inverting amplifier circuit is infinite since no current flows into the inverting input. Output impedance is zero since output voltage is ideally independent of output current. Closed loop gain is $1 + \frac{R_O}{R_I}$, hence can be any desired value above unity. Such circuits are widely used in control and instrumentation where non-inverting gain is required.

3.6 Inverting Amplifier

The inverting amplifier appears in Figure 3.3(b). This circuit and its many variations form the bulk of commonly used operational amplifier circuitry. Single ended input and output versions were first used, and they became the basis of analog computation. Today's modern differential input amplifier is used as an inverting amplifier by grounding the non-inverting input and applying the input signal to the inverting input terminal.

Since the amplifier draws no input current and the input voltage approaches zero when the feedback loop is closed (the two summing point restraints), we may write:

$$\frac{E_I}{R_I} + \frac{E_O}{R_O} = 0$$

Hence

$$\frac{E_O}{E_I} = -\frac{R_O}{R_I}$$

Input impedance to this circuit is not infinite as in the two previous circuits, the inverting input is at ground potential so the driving source effectively "sees" R_{I} as the input impedance. Output impedance is zero as in the two previous circuits. Closed loop gain of this circuit is $-\dfrac{R_{\mathrm{O}}}{R_{\mathrm{I}}}$.

3.7 Intuitive Analysis Techniques

The popularity of the inverting amplifier has been mentioned already. *In control and instrumentation applications, its practical value lies in the ease with which desired input impedance and gain values can be tailored to fit the requirements of the associated circuitry*[5]. Its utility is reflected in the variety of intuitive devices that are commonly used to simplify its analysis.

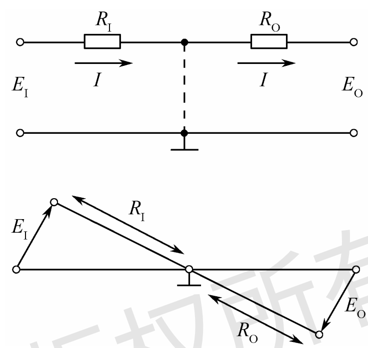


Figure 3.7 Intuitive devices for analysis of circuits based on the inverting amplifier

If we draw the summing point, the inverting input and output of the inverting amplifier as in Figure 3.7, the dotted line serves as a reminder that the inverting input is at ground potential but conducts no current to ground. The output can supply any needed current, and analysis quickly becomes rote. Another such device uses the action of a lever to show as vectors the voltage relations that exist using the inverting input as the fulcrum.

Words & Expressions

buffer	['bʌfə]	*n.*	缓冲器
by-product	['baiprɔdʌkt]	*n.*	副产品，出乎意料的结果
circuitry	['sə:kitri]	*n.*	电路，线路
formula	['fɔ:mjulə]	*n.*	公式，规则
fulcrum	['fʌlkrəm]	*n.*	杠杆的支点
ground	[graund]	*v.*	使接地　*n.* 接地，地线
impedance	[im'pi:dəns]	*n.*	阻抗
infinitesimal	[in,finə'tesiməl]	*adv.*	无穷小，极小，无限小
inherent	[in'hiərənt]	*adj.*	固有的，内在的，与生俱来的
intuitive	[in'tju(:)itiv]	*adj.*	直觉的
investigate	[in'vestigeit]	*v.*	调查，研究
latch	[lætʃ]	*v.*	闭锁（*n.* 锁存器）
lever	['li:və, 'levə]	*n.*	杆，杠杆
mythical	['miθikəl]	*adj.*	神话的，虚构的
offset	['ɔ:fset]	*n.*	偏移量，抵消

polarity	[pəu'læriti]	*n.*	极性
potential	[pə'tenʃəl]	*n.*	电势，电位
restraint	[ris'treint]	*n.*	约束（条件）
rote	[rəut]	*n.*	死记硬背，机械的做法，生搬硬套
saturation	[ˌsætʃə'reiʃən]	*n.*	饱和（状态），饱和度
stability	[stə'biliti]	*n.*	稳定性
unity	['juːniti]	*n.*	统一，一致
variation	[ˌvɛəri'eiʃən]	*n.*	变更，变化，变异，变种
closed loop gain			闭环增益
differential input amplifier			差分放大器
inverting input			反相输入端
ideal operation amplifier			理想运放（*abbr*. ideal op amp）
phase shift			相移
voltage follower			射极跟随器

Notes

[1] Frequency response will be flat and bandwidth infinite because AC will be simply a rapidly varying DC level to the ideal amplifier.　bandwidth 后省略 will be。

对于理想放大器而言，交流只不过是快速变化的直流，所以其频率响应是一条水平直线，带宽无限。

[2] An important by-product of these properties of the ideal operational amplifier is that the summing point, the inverting input, will conduct no current to the amplifier.　the inverting input 是 the summing point 的同位语，表示相同的意思，翻译时可以省略。

由理想运放的特性可以得出一个很重要的结论，即放大器的反相输入端无电流流过。

[3] Though the ideal model may seem a bit remote from reality — with infinite gain, bandwidth, etc., it should be realized that the closed loop gain relations which will be derived in this section are directly applicable to real circuits — to within a few tenths of a percent in most cases.　with 结构对从句中的主语 the ideal model 予以补充说明，a few tenth of 表示“十分之几”的意思。

尽管理想模型与实际电路相差甚远，比如，理想运放具有无限带宽和无穷大的增益，但需要明白的是，本节利用理想运放推导出的闭环增益公式可直接应用于实际电路，多数情况下两者相差仅为千分之几。

[4] If a small voltage, measured at the inverting input with respect to the non-inverting input, is assumed to exist, the amplifier output voltage will be of opposite polarity and can always increase in value (with infinite output available) until the voltage between the inputs becomes infinitesimally small.　句中的 opposite 是相对于从句中的主语 a small voltage 而言。

假如在反相输入端测得与同相输入端之间存在一小电压，则放大器的输出电压与该输入电压极性相反且其数值会一直增大（可以是无穷大），直至输入电压变为无穷小为止。

[5] In control and instrumentation applications, its practical value lies in the ease with which desired input impedance and gain values can be tailored to fit the requirements of the associated circuitry.　lie in

表示原因，the ease with which… 中的 which 指代 ease, with ease 等效于 easily。

在控制与仪器仪表应用中，反相放大器的实用价值在于，利用它可以很容易地调整所需的输入阻抗和增益值以适应相关电路的要求。

Grammar

数量的表示

科技文献中经常有大量数词出现。数词表示事物的数量或数目，其含义十分严格，理解或翻译时的疏忽和差错可能会引起严重后果。由于英语与汉语在数量表达上差别较大，对待数词或数量的表达要特别仔细。

1. 数字的表示

科技文章中，数字频繁出现，用阿拉伯数字比用单词陈述更有利。但出现下述情形时须遵循约定俗成的规则：用单词表示不定数量或近似值；句首不用阿拉伯数字，一般用英语的单词，句末要尽量避免用阿拉伯数字；两数连用时，分别用单词和阿拉伯数字表示，习惯上将短的用单词写出；遇到分数时，可用带连字符号的单词表示，等等。举例如下。

[例 1] Phase shift is 180°.

（输出电压与输入电压的）相移为 180°。

[例 2] The gain of the voltage follower with the feedback loop closed (closed loop gain) is *unity*.

电压跟随器的闭环增益为 1。

2. 不确定数字的表示（见表 3.1）

① 大约、左右，常用 about/some/approximately/of the order of/more or less/ or so 等词；
② 多于，常用 over, above, more than, in excess of 等词；
③ 少于，常用 less than, under, below, close to 等词；
④ 以复数形式表示，如 tens/dozens/scores/hundreds/thousands of, thousands upon thousands of 等。

表 3.1　某些表示数量的词和词组

修饰可数名词	修饰不可数名词	修饰可数或不可数名词
few（几乎没有）	little	some（一些） a lot of plenty of
not many（不多）	not much	
a few（少数几个）	a little	
a great many (很多)	a great deal	
a { small / moderate / certain / negligible / large / great / considerable / substantial } { number of（可数）/ amount of（不可数）}	少量的 适量的 一定量的 不多的 多种 多种 很多 很多	a { certain / moderate / large } { percentage / proportion / part / quantity } of

[例 3] The result indicated that actual error probabilities were *of the order of* 1 percent.

结果表明实际误差概率约为 1%。

[例 4] The microwave communication channel has a very large bandwidth and will accommodate *thousands of* telephone conversations or *dozens of* television channels at once.

微波通信信道带宽很宽，可同时容纳几千个电话通话或几十个电视信道。

3. 习惯短语

[例 5] Economies associated with computer-on-a-chip have resulted in the availability of micro-computer systems with the functionality and performance of minicomputer systems costing *two orders of magnitude* more only a decade ago.

随着单片机的发展，目前微机系统所拥有的功能及性能可以媲美 10 年前的小型机，而价格却下降了两个数量级以上。

[例 6] The resistance of a given section of an electric circuit is equal to *the ratio of* its voltage *to* the current through this section of the circuit.

电路中某部分的电阻等于它两端的电压与流过该部分电路的电流的比值。

[例 7] It is found that the heat energy developed in any conductor *is proportional to* (*is in proportion to*) the square of the current, the resistance of the conductor and the time.

人们发现，导体中所产生的热能与电流的平方、导体的电阻值和时间成正比。

4. 倍数增减

倍数的增减在汉语与英语中的表述有较大的差异，在汉语中增加可以是倍数分数，而减少则只能是分数，而倍数的增减又常常涉及是否将基数计入其中，所以需要仔细领会。

① *n* times as ＋ 形容词 / 副词 ＋ as
n times ＋ 名词或 that…
n times ＋ 比较级 ＋than
是……的 *n* 倍，增加了(n–1)倍，比……大（高，长等）(n–1)倍
是……的 1/*n*，比……小(n–1)/*n*

[例 8] The power density from a transmitter measured 1m from an isotropic antenna is *four times as large as* the power density measured 2m from the same antenna.

在离各向同性天线 1 米远的距离测量的功率密度是离其 2 米远测量的功率密度的 4 倍。

② 倍数增减的词+ *n* times / to *n* times / *n*-fold / by a factor of *n*，译为“增至 *n* 倍，或增加了 *n*–1 倍，减少到 1/*n*，或减少了(n–1)/*n*”

[例 9] The production of ICs has been *increased to three times* as compared with last year.

集成电路的产量比去年增加了两倍。

[例 10] The switching time of the new-type transistor is *shortened 3 times*.

新型晶体管的开关时间缩短了 2/3（即缩短到原来的 1/3）。

[例 11] An algorithm is given that *reduces* the number of multiplications by almost *a factor of two* but increases the number of additions.

给出的新算法中，乘法数减少了一半而加法数却有所增加。

③ a/an *n* times / *n*-fold + 表示增减的名词，应译为：增加了 *n*– 1 倍（或增至 *n* 倍），或减少到 1/*n*，减少了(n–1)/*n*

[例 12] The principal advantage of the new type mobile is *a four-fold reduction* in power dissipation.

新手机的主要优点是功耗降低了 3/4。

④ 增加一倍、两倍、三倍分别用 double（是……的两倍），treble（是……的三倍），quadruple（是……的 4 倍）表示，再往上就用 *n* times / *n*-fold。减半常用 break/cut/split in half/into halves, decrease one-half, shorten … two times 等表示。

[例 13] *Reducing* the data rate *by one-half* will *double* the duration of each symbol interval.

数据率减半，将使每个符号间隔的持续时间增加一倍。

Exercises

1. Choose the best answer for each of the following questions.

（1）An ideal operational amplifier has ________.

a. infinite gain, infinite bandwidth, infinite input impedance, zero output impedance

b. infinite gain, infinite bandwidth, zero input impedance, infinite output impedance

c. infinite gain, zero bandwidth, infinite input impedance, infinite output impedance

d. zero gain, infinite bandwidth, infinite input impedance, infinite output impedance

（2）One of the summing constraints is ________.

a. the inverting input conducts no current to the amplifier

b. no current flows into either input terminal of the ideal op amp

c. the inverting input is connected to the noninverting input

d. the voltage between the two input terminals is non-zero

（3）For an inverting amplifier, ________.

a. the output voltage is out of phase with the input voltage by 180°

b. the phase difference between the input and output voltage is 90°

c. the source signal is applied to the non-inverting input terminal

d. there exits a positive feedback between the output and the input terminal

（4）Which is not true about the voltage follower? ________.

a. It is of no use because the output voltage is the same as that of the input

b. It is a special case of the noninverting amplifier

c. It can be used to isolate circuits and prevent undesired interaction

d. It may allow sources with low current capability to drive a heavy load

（5）For a voltage follower, ________.

a. the input impedance is very small

b. the output impedance is large

c. it has unity gain

d. there is current flowing through the feedback element

2. Match each term in column A with its explanation in Column B.

Column A	**Column B**
（1）Output Peak Current	a. the ratio between the input and output signals

(2) Output Voltage Swing	b. the minimum value of the small-signal gain bandwidth product (in Hz) of a voltage amplifier at specified frequency and at reference conditions
(3) Open-Loop Gain (A_V)	c. the maximum dissipation that the amplifier can safely support
(4) Supply Voltage Range	d. the maximum rate of change of the output signal in response to a step input
(5) Power Dissipation	e. the maximum current that can be delivered in the output of the amplifier
(6) Gain Bandwidth Product	f. the total impedance as seen from the input terminals
(7) Common Mode Rejection Ratio	g. the maximum (positive and negative) output voltage
(8) Slew Rate	h. a measure of the ability of the amplifier to produce zero output voltage when the two inputs have the same voltage
(9) Input Impedance	i. the range of voltages that can be applied to an amplifier
(10) Quiescent Current	j. the DC current required by the inputs of the amplifier to properly drive the first stage
(11) Input Offset Voltage (V_{OS})	k. the current produced by the amplifier when in normal operation
(12) Input Bias Current (I_{BIAS})	l. a small DC output voltage when 0V is applied to the input

3. Translate into Chinese.

(1) Operational amplifiers can have either a closed-loop operation or an open-loop operation. The operation (closed-loop or open-loop) is determined by whether or not feedback is used.

(2) In $A_{vf} = \dfrac{A_v}{1 + A_v F}$, A_{vf} is called the close loop gain of the op amp circuit, whereas A_v is the open loop gain. The product $A_v F$ is called the loop gain. This is the gain a signal would see starting at the inverting input and traveling in a clockwise loop through the op amp and the feedback network.

(3) Although it is useful and easy to treat the op amp as a black box with a perfect input/output characteristic, it is important to understand the inner workings, so that one can deal with problems that will arise due to internal parasitic capacitances, etc.

(4) When feedback is used around an operational amplifier, the closed loop gain of the circuit is determined by a ratio involving the input and feedback impedances used.

(5) In the case of the ideal operational amplifier, circuit analysis was simplified by the ideal summing point restraints of zero voltage and zero current at the inverting input.

(6) The operational amplifier is an extremely efficient and versatile device. Its applications span the broad electronic industry filling requirements for signal conditioning, special transfer functions, analog instrumentation, analog computation, and special systems design.

4. Read the following article, pay attention to the expressions in italics.

CITING strong communications and consumer chip demand, Taiwan Semiconductor Manufacturing Co (TSMC) reported record sales and *a six-fold increase* in net profit for *the fourth* quarter of last year. It also said it plans to increase capital expenditure *more than* 60 *percent* this year to US$2 billion — indicating that an industry recovery is well under way.

The world's largest made-to-order semiconductor manufacturer, or foundry, recorded *a fourth quarter* net profit of NT$16.02 billion (US$815.3 million) — *a* 526.9 *percent increase* year on year. Revenue was NT$57.78 billion, *up* 5.3 *percent* from *the third quarter* and 40.4 *percent* from a year earlier.

TSMC closed 2003 with sales of NT$201.9 billion, *an increase of* 25.6 *percent* from 2002, and income of NT$47.266 billion, *an increase of* 118.7 *percent* from the previous year.

Growth was driven largely by communications products, said TSMC chairman Morris Chang. Communication chip sales *grew* from *about* 35 *percent* of total sales in the fourth quarter of 2002 to 42 *percent* of total sales by the fourth quarter of last year.

According to Tony Massimini, chief of technology research at US based Semico Research Corp: "The recovery will be driven by consumer and communications markets." Cellular phone handset sales are forecast to *grow about 10 percent* to *more than half a billion* units this year, and the chip content in many consumer products is *doubling* or *tripling* due to increased functions.

The foundry industry as a whole will grow *as much as* 43 *percent* this year, according to some forecasts. Such anticipated growth is leading to increased spending.

Mr Chang said yesterday TSMC will spend US$2 billion to build capacity, and US$1.5 billion of this will be spent on building 12-inch capacity.

TSMC's rival United Microelectronics Corp is expected to increase capital expenditure *more than four-fold* this year.

Reading Material

Digital System Design Hierarchy

Digital systems may be designed and studied at many different levels of abstraction, ranging from a purely behavioral model, in which no hardware details are specified, down to the physical level, in which only structures of physical materials are specified. Several levels of design abstraction are listed in Table1.

Table 1 Hierarchy of system design abstraction

Design Level	Level of Abstraction	Amount of Detail	Type of Model
System	Highest	Lowest	Behavioral
Register	—	—	Behavioral/structural
Gate	—	—	Structural
Transistor	—	—	Structural
Physical	Lowest	Highest	Structural

1. The System and Register Levels

At its highest level, a digital system can be viewed as one or more interacting functional modules. The behavior of each module is described without specifying implementation details. For example, a desktop computer viewed at the system level comprises a microprocessor, memory modules, and control circuits for the monitor, keyboard, printer, and other peripheral devices.

At the register level, a digital system is viewed as a collection of elements called registers that store information, interconnected in some fashion by signal lines. Information is processed by the system by transferring it between registers along these signal lines. In some cases the information is transformed during these register transfers by routing it through one or more functional modules. Figures 1(a) and (b) illustrate the system- and register-level models of a digital system that computes the sum of a sequence of binary numbers, supplied one at a time as inputs to the system. At the system level, all that is known is the basic function of the system, which is to compute:

$$\text{Total} = \sum_{i=1}^{N} \text{Input}_i$$

At the register level, as in Figure 1(b), it is seen that the system comprises a storage register A and an adder circuit. The Total is computed by first clearing register A, using control signal *Clear*, and then adding each input number, Input_i, to the contents of register A, replacing the contents of register A with the new sum, using control signal store. Hence, the sum of a list of numbers is computed by performing the following register transfers in the proper sequence.

$$\text{Clear:} \quad A \leftarrow 0$$

$$\text{Store:} \quad A \leftarrow A + \text{input}$$

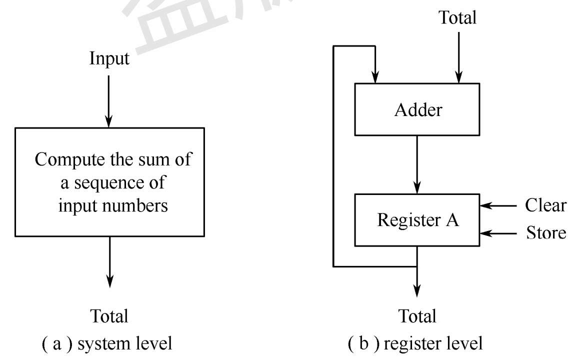


Figure 1 Models of a digital system that adds lists of numbers

2. The Gate Level

At its lowest level, the behavior of a digital system is specified as a set of logic equations from switching algebra that can be realized in hardware by logic circuits. The smallest logical unit of digital hardware is called a gate. Gates are switching elements that implement the fundamental operators of switching algebra. Logic equations are realized in hardware by interconnecting gates to form combinational logic circuits, as illustrated in Figure 2. Note that the circuit has six gates. The inputs in this example are labeled $x_1, \cdots, x_5$, and the output $f\left(x_1, \cdots, x_5\right)$ is a function only of the present value of the input signals. Hence, a distinguishing feature of the combinational logic circuit is that it possesses no memory of previous input signals. The analysis and design of combinational logic circuits consume a major portion of this text.

All digital computers contain memory devices called registers that serve as temporary stores for information. These registers and certain parts of the control unit are called sequential logic circuits. A sequential logic circuit is, in general, a combinational logic circuit with memory, as modeled in Figure 3. Unlike combinational logic circuits, the outputs of a sequential logic circuit are functions of not only the present value of the input signals, but also depend on the past history of inputs, as reflected by the information stored in the registers. Sequential logic circuit analysis and design comprise the second focal point of this text. Only after readers have mastered the fundamentals of combinational and sequential circuits can they proceed with the design and construction of digital systems hardware.

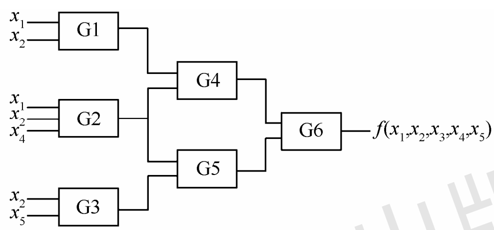


Figure 2　A combinational logic circuit with six gates

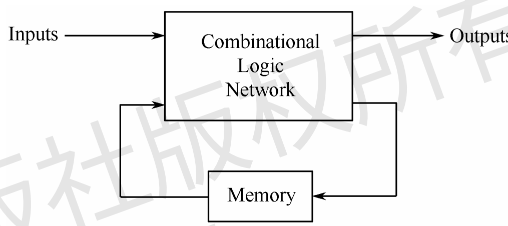


Figure 3　Sequential logic circuit

3. Transistor and Physical Design Levels

Combinational and sequential logic circuits completely define the logical behavior of a digital system. Ultimately, each logic gate must be realized by a lower-level transistor circuit, which in turn is realized by combining various semiconductor and other materials. The technologies used to construct gates and other logic elements have evolved from mechanical devices to relays to electron tubes to discrete transistors to integrated circuits. Modern computers and application-specific digital systems are usually built of integrated circuits that are arranged to realize the registers and control circuits necessary to implement the computer's instruction set or the system's functions.

An integrated circuit (IC) contains multiple logic elements. The number of gates or gate equivalents per IC determines the scale of integration. Small scale integration (SSI) refers to ICs with 1 to 10 gates, medium scale integration (MSI) corresponds to 10 to 100-gate ICs, large scale integration (LSI) 100 to 10,000 gates, and very large scale integration (VLSI) to ICs with more than 100000 gates.

It is beyond the scope of this text to consider transistor and physical-level design of logic gates. However, it is important to have a basic understanding of various electrical and physical properties of different gate circuits so that the logical operation, performance, cost, and other parameters of a digital system design may be evaluated.

4. Electronic technologies

Numerous families of electronic technologies have been developed to provide characteristics such as speed, power consumption, packaging density, functionality, and cost that hardware designers prefer. Usually, it is impossible to provide all the desired characteristics in one family. Hence, there is an ongoing quest for improvements in proven technologies or the development of new technologies. Table 2 and 3 list the most significant technologies and corresponding characteristics that have been used since the beginning of the transistor era.

The packaging of logic gates and other logic elements has changed significantly over the years. Early electronic logic elements were typically constructed from large electron tubes, discrete resistors, and capacitors, were mounted on an aluminum chassis, and were interconnected with copper wire. Tube technology advances resulted in reduced sizes, and printed circuit boards replaced the wires. Later, discrete transistors replaced the tubes, but the resistors, capacitors, and printed circuit boards remained in use, although their sizes were smaller: the advent of the integrated circuit in the early 1960s produced further reduction in the size of printed circuit boards and other passive elements.

Integrated circuits can be manufactured in standard, semicustom, and custom forms. Standard ICs provide the pans necessary to build systems for most applications. However, some applications may require semicustom or custom circuits to meet special functions, lower cost, or smaller size requirements. Custom circuits are manufactured to the exact requirements of a specific customer. On the other hand, semicustom circuits are programmed to satisfy a customer's need. The term *application-specific* integrated circuits (ASICs) is often used to describe semicustom devices.

Table 2　Important electronic technologies

Technology	Device type
Resistor-transistor logic (RTL)	Bipolar junction
Diode-transistor logic (DTL)	Bipolar junction
Transistor-transistor logic (TTL)	Bipolar junction
Emitter-coupled logic (ECL)	Bipolar junction
Positive metal oxide semiconductor (PMOS)	MOSFET
Negative metal oxide semiconductor (NMOS)	MOSFET
Complementary metal oxide semiconductor (CMOS)	MOSFET
Gallium Arsenide (GaAs)	MESFET

Table 3　Characteristics of electronic technology families

Technology	Power consumption	Speed	Packaging
RTL	High	Low	Discrete
DTL	High	Low	Discrete, SSI
TTL	Medium	Medium	SSI, MSI
ECL	High	High	SSI, MSI, LSI
pMOS	Medium	Low	MSI, LSI
nMOS	Medium	Medium	MSI, LSI, VLSI
CMOS	Low	Medium	SSI, MSI, LSI, VLSI
GaAs	High	High	SSI, MSI, LSI

New words

capacitor	电容器
diode	二极管
family	族，一群相似的事物
gate	（逻辑）门
operator	算子
passive	无源的
register	寄存器
route	（经由某一路线）发送
semicustom	半定制
combinational logic circuit	组合逻辑电路
peripheral devices	外设
printed circuit boards	印制电路板
sequential logic circuit	时序逻辑电路
ASICs (application-specific integrated circuits)	专用集成电路
GaAs (allium arsenide)	砷化镓
IC (integrated circuits)	集成电路
MOS (metal oxide semiconductor)	金属氧化物半导体

Questions

1. What is the difference between the combinational and sequential logic circuits?
2. Give an example of which can be implemented using the combinational logic circuit.
3. How do you determine whether a system belongs to a VLSI or a MSI?
4. Describe the three forms used in the manufacturing of integrated circuits.

Unit 4

Text

Signals, Linear Systems, and Convolution

Characterizing the complete input-output properties of a system by exhaustive measurement is usually impossible. *Instead, we must find some ways of making a finite number of measurements that allow us to infer how the system will respond to other inputs that we have not yet measured*[1]. We can only do this for certain kinds of systems with certain properties. If we have the right kind of system, we can save a lot of time and energy by using the appropriate theory about the system's responsiveness. Linear systems theory is a good time-saving theory for linear systems which obey certain rules. Not all systems are linear, but many important ones are. When a system qualifies as a linear system, it is possible to use the responses to a small set of inputs to predict the response to any possible input. This can save the scientist enormous amounts of work, and makes it possible to characterize the system completely.

4.1 Continuous-Time and Discrete-Time Signals

Anything that bears information can be considered a signal. *The type of signal which depends on one independent variable, namely, time is called the one-dimensional signal. It can be represented by* $x(t)$[2]. For every t, the signal is required to assume a unique value; otherwise the signal is not well defined. Consider the temperature $x(t)$. It is defined at every time instant and is called a continuous-time (CT) signal. A CT signal is also called an analog signal because its waveform is often analogous to that of the physical variable. Other examples of analog signals are the waveform of the potential at household electric outlet, and the electrocardiogram (EKG) waveforms.

A signal is called a discrete-time (DT) signal if the signal is defined at discrete time instant. Most discrete-time signals arise from sampling continuous-time signals. For example, if the temperature $x(t)$ is measured and recorded every hour, then the resulting signal will be denoted by

$$x[n] := x(nT) = x(t)\big|_{t=nT}$$

where T is called the sampling period and n is an integer ranging from $-\infty$ to ∞ called the time index. We call $x[n]$ the sampled signal or sampled sequence of $x(t)$.

The vast majority of signals encountered in practice are continuous-time or analog signal. It is often mathematically convenient to work with continuous-time signals. But in practice, you usually end up with discrete-time sequences because: (1) discrete-time samples are the only things that can be measured and recorded when doing a real experiment; and (2) finite-length, discrete-time sequences are the only things that can be stored and computed with computers.

In what follows, we will express most of the mathematics in the continuous-time domain. But the examples will, by necessity, use discrete-time sequences.

Pulse and impulse signals The unit impulse signal, written $\delta(t)$, is one at $t=0$, and zero everywhere else:

$$\delta(t)=\begin{cases}1 & \text{if} \quad t=0 \\ 0 & \text{otherwise}\end{cases}$$

Unit step signal The unit step signal, written $u(t)$, is zero for all times less than zero, and 1 for all times greater than or equal to zero:

$$u(t)=\begin{cases}0 & \text{if} \quad t<0 \\ 1 & \text{if} \quad t \geqslant 0\end{cases}$$

Summation and integration An integral is the limiting case of a summation:

$$\int_{-\infty}^{\infty} x(t) \mathrm{d} t=\lim _{\Delta \rightarrow 0} \sum_{k=-\infty}^{\infty} x(k \Delta) \Delta$$

For example, the step signal can be obtained as an integral of the impulse:

$$u(t)=\int_{-\infty}^{t} \delta(s) \mathrm{d} s$$

Up to $s<0$ the sum will be 0 since all the values of $\delta(t)$ for negative s are 0. At $t=0$ the cumulative sum jumps to 1 since $\delta(t)=1$. And the cumulative sum stays at 1 for all values of t greater than 0 since all the rest of the values of $\delta(t)$ are 0 again.

Representing signals with impulses Any signal can be expressed as a sum of scaled and shifted unit impulses. We begin with the pulse or "staircase" approximation $\tilde{x}(t)$ to a continuous signal $x(t)$, as illustrated in Figure 4.1. Conceptually, this is trivial: for each discrete sample of the original signal, we make a pulse signal. Then we add up all these pulse signals to make up the approximate signal. Each of these pulse signals can in turn be represented as a standard pulse scaled by the appropriate value and shifted to the appropriate place. In mathematical notation:

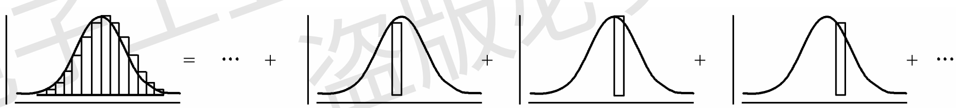


Figure 4.1 Staircase approximation to a continuous-time signal

$$\tilde{x}(t)=\sum_{k=-\infty}^{\infty} x(k \Delta) \delta_{\Delta}(t-k \Delta) \Delta$$

As we let Δ approach zero, the approximation $\tilde{x}(t)$ becomes better and better, and the approximation in the limit equals $x(t)$. Therefore,

$$x(t)=\lim _{\Delta \rightarrow 0} \sum_{k=-\infty}^{\infty} x(k \Delta) \delta_{\Delta}(t-k \Delta) \Delta$$

Also, as $\Delta \rightarrow 0$, the summation approaches an integral, and the pulse approaches the unit impulse:

$$x(t)=\int_{-\infty}^{\infty} x(s) \delta(t-s) \mathrm{d} s \tag{1}$$

In other words, we can represent any signal as an infinite sum of shifted and scaled unit impulses. A digital compact disc, for example, stores whole complex pieces of music as lots of simple numbers representing very short impulses, and then the CD player adds all the impulses back together one after another to recreate the complex musical waveform.

4.2 Linear Systems

A system or transform maps an input signal $x(t)$ into an output signal $y(t)$:

$$y(t) = T[x(t)]$$

where T denotes the transform, a function from input signals to output signals.

Systems come in a wide variety of types. One important class is known as linear systems. To see whether a system is linear, we need to test whether it obeys certain rules that all linear systems obey. The two basic tests of linearity are homogeneity and additivity.

Homogeneity As we increase the strength of the input to a linear system, say we double it, then we predict that the output function will also be doubled. For example, if the current injected to a passive neural membrane is doubled, the resulting membrane potential fluctuations will double as well. This is called the *scalar rule* or sometimes the *homogeneity* of linear systems.

Additivity Suppose we measure how the membrane potential fluctuates over time in response to a complicated time-series of injected current $x_1(t)$. Next, we present a second (different) complicated time-series $x_2(t)$. The second stimulus also generates fluctuations in the membrane potential which we measure and write down. Then, we present the sum of the two currents $x_1(t)+x_2(t)$ and see what happens. Since the system is linear, the measured membrane potential fluctuations will be just the sum of the fluctuations to each of the two currents presented separately.

Superposition Systems that satisfy both homogeneity and additivity are considered to be linear systems. These two rules, taken together, are often referred to as the principle of superposition. Mathematically, the principle of superposition is expressed as:

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) \qquad (2)$$

Homogeneity is a special case in which one of the signals is absent. Additivity is a special case in which $\alpha + \beta = 1$.

Shift-invariance Suppose that we inject a pulse of current and measure the membrane potential fluctuations. Then we stimulate again with a similar pulse at a different point in time, and again we measure the membrane potential fluctuations. If we haven't damaged the membrane with the first impulse then we should expect that the response to the second pulse will be the same as the response to the first pulse. The only difference between them will be that the second pulse has occurred later in time, that is, it is shifted in time. *When the responses to the identical stimulus presented shifted in time are the same, except for the corresponding shift in time, then we have a special kind of linear system called a shift-invariant linear system*[3]. Just as not all systems are linear, not all linear systems are shift-invariant.

In mathematical language, a system T is shift-invariant if and only if

$$y(t) = T[x(t)] \quad \textbf{implies} \quad y(t-s) = T[x(t-s)] \qquad (3)$$

To get a better feeling for linearity, think about a technician trying to determine if an electronic device is linear. The technician would attach a sine wave generator to the input of the device, and an oscilloscope to the output. With a sine wave input, the technician would look to see if the output is also a sine wave. For example, the output cannot be clipped on the top or bottom, the top half cannot look different from the bottom half, there must be no distortion where the signal crosses zero, etc.

Next, the technician would vary the amplitude of the input and observe the effect on the output signal. If the system is linear, the amplitude of the output must track the amplitude of the input. Lastly, the technician would vary the input signal's frequency, and verify that the output signal's frequency changes accordingly. As the frequency is changed, there will likely be amplitude and phase changes seen in the output, but these are perfectly permissible in a linear system. At some frequencies, the output may even be *zero*, that is, a sinusoid with zero amplitude. If the technician sees all these things, he will conclude that the system is linear. *While this conclusion is not a rigorous mathematical proof, the level of confidence is justifiably high*[4].

Convolution Homogeneity, additivity, and shift invariance may, at first, sound a bit abstract but they are very useful. To characterize a shift-invariant linear system, we need to measure only one thing: the way the system responds to a unit impulse. This response is called *the impulse response function* of the system. Once we've measured this function, we can (in principle) predict how the system will respond to any other possible stimulus. In the following, we show that the response of a shift-invariant linear system can be written as a sum of scaled and shifted copies of the system's impulse response function.

The convolution integral Begin by using Eq.(1) to replace the input signal $x(t)$ by its representation in terms of impulses:

$$y = T\left[x(t)\right] = T\left[\int_{-\infty}^{\infty} x(s)\delta(t-s)\mathrm{d}s\right] = T\left[\lim_{\varDelta\to 0}\sum_{k=-\infty}^{\infty} x(k\varDelta)\delta_{\varDelta}(t-k\varDelta)\varDelta\right]$$

Using additivity,

$$y(t) = \lim_{\varDelta\to 0}\sum_{k=-\infty}^{\infty} T\left[x(k\varDelta)\delta_{\varDelta}(t-k\varDelta)\varDelta\right]$$

Taking the limit,

$$y(t) = \int_{-\infty}^{\infty} T\left[x(s)\delta(t-s)\mathrm{d}s\right]$$

Using homogeneity,

$$y(t) = \int_{-\infty}^{\infty} x(s)T\left[\delta(t-s)\right]\mathrm{d}s$$

Now let $h(t)$ be the response of T to the unshifted unit impulse, i.e., $h(t) = T[\delta(t)]$. Then by using shift-invariance,

$$y(t) = \int_{-\infty}^{\infty} x(s)h(t-s)\mathrm{d}s \tag{4}$$

Notice what this last equation means. *For any shift-invariant linear system T, once we know its impulse response h*(*t*) (*that is, its response to a unit impulse*), *we can forget about T entirely, and just add up scaled and shifted copies of h*(*t*) *to calculate the response of T to any input whatsoever*[5]. Thus any shift-invariant linear system is completely characterized by its impulse response $h(t)$.

The way of combining two signals specified by Eq.(4) is known as convolution. It is such a widespread and useful formula that it has its own shorthand notation, *. For any two signals x and y, there will be another signal z obtained by convolving x with y,

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

Convolution as a series of weighted sums *While superposition and convolution may sound a little abstract, there is an equivalent statement that will make it concrete: a system is a shift-invariant, linear system if and only if the responses are a weighted sum of the inputs*[6]. The choice of weighting function determines the behavior of the system. Not surprisingly, the weighting function is very closely related to the impulse response of the system. In particular, the impulse response and the weighting function are time-reversed copies of one another.

Properties of convolution The following things are true for convolution in general, as you should be able to verify for yourself with some algebraic manipulation:

$x*y=y*x$	commutative
$(x*y)*z=x*(y*z)$	associative
$(x*z)+(y*z)=(x+y)*z$	distributive

Words & Expressions

additivity	[ˌædi'tiviti]	*n.*	叠加性
associative	[ə'səuʃjətiv]	*adj.*	结合的
commutative	[kə'mjuːtətiv]	*adj.*	交换的
cumulative	['kjuːmjulətiv]	*adj.*	累加的
conceptual	[kən'septjuəl]	*adj.*	概念的
convolution	[ˌkɔnvə'ljuːʃən]	*n.*	卷积
distortion	[dis'tɔːʃən]	*n.*	扭曲，变形，失真
distributive	[dis'tribjutiv]	*adj.*	分配的
electrocardiogram	[ilektrəu'kɑːdiəugræm]	*n.*	心电图（ECG）
enormous	[i'nɔːməs]	*adj.*	巨大的，庞大的
fluctuation	[ˌflʌktju'eiʃən]	*n.*	波动，起伏
generator	['dʒenəreitə]	*n.*	发生器，产生器
homogeneity	[ˌhɔməudʒe'niːiti]	*n.*	齐次性
identical	[ai'dentikəl]	*adj.*	相同的
map	[mæp]	*n.*	映射
membrane	['membrein]	*n.*	膜，隔膜
negative	['negətiv]	*adj.*	负的
oscilloscope	[ɔ'siləskəup]	*n.*	示波器
permissible	[pə(ː)'misəbl]	*adj.*	可允许的
rigorous	['rigərəs]	*adj.*	严格的
sampling	['sɑːmpliŋ]	*n.*	采样
stimulate	['stimjuleit]	*v.*	刺激
superposition	[ˌsjuːpəpə'ziʃən]	*n.*	重叠，叠加
unique	[juː'niːk]	*adj.*	唯一的
waveform	['weivfɔːm]	*n.*	波形
analog signal			模拟信号

one-dimensional	一维的
shift-invariance	时不变特性

Notes

[1] Instead, we must find some ways of making a finite number of measurements that allow us to infer how the system will respond to other inputs that we have not yet measured.　定语从句 that allows us…的先行词为 ways，而非 measurements.

因此，必须找到一些方法，通过一定数量的测量就能推断出系统对其他未经测量的输入如何响应。

[2] The type of signal which depends on one independent variable, namely, time is called the one-dimensional signal. It can be represented by $x(t)$.　time 用于具体说明 one independent variable.

称取决于某一独立变量，即时间 t 的这类信号为一维信号，用 $x(t)$表示。

[3] When the responses to the identical stimulus presented shifted in time are the same, except for the corresponding shift in time, then we have a special kind of linear system called a shift-invariant linear system.　本句中 when 引导的状语从句实际上表示条件，而非时间。

如果系统对经时移后的相同激励产生具有相应时移的相同响应，我们就得到一种特殊的线性系统，称为“线性时不变系统”。

[4] While this conclusion is not a rigorous mathematical proof, the level of confidence is justifiably high.　副词 justifiably 对应的动词是 justify，表示“证明……正当/有理由”之意。

尽管该结论并非严格的数学证明，但有理由相信其可信度很高。

[5] For any shift-invariant linear system T, once we know its impulse response $h(t)$ (that is, its response to a unit impulse), we can forget about T entirely, and just add up scaled and shifted copies of $h(t)$ to calculate the response of T to any input whatsoever.　代词 whatsoever 为 input 的后置定语，表示“无论什么”之意。

对任意线性时不变系统 T，只要已知冲激响应 $h(t)$（即系统对单位冲激函数的响应），就可以不用考虑 T，将经过缩放和位移变换的各个 $h(t)$相加，即可得到系统 T 对任意输入的响应。

[6] While superposition and convolution may sound a little abstract, there is an equivalent statement that will make it concrete: a system is a shift-invariant, linear system if and only if the responses are a weighted sum of the inputs.　本句中 sound 为不及物动词表示“听起来”，if and only if 说明充分必要条件，常译为“当且仅当”，也常缩写为 iff。

叠加与卷积听起来可能有些抽象，但具体可等效解释为：当且仅当系统的响应可表示为输入信号的加权和时，该系统才是线性时不变系统。

Grammar

常用介词及其用法

介词在科技英语中的用法较多，它常与名词、动词或形容词等搭配构成介词短语，也可构

成短语介词，其作用相当于单个介词。熟悉科技英语中常用的介词与其他词类的习惯搭配是正确理解词意的一种重要手段。

1. 介词与动词、名词、形容词搭配

① 与动词（绝大部分及物）构成短语动词，要求后接宾语。例如：apply to, arise from, consist of, deal with, depend on/upon, relate to, result in, differ from, follow from, lead to, lie in, vary with 等。

[例 1] A complex variable *s is composed of* a real part alpha and an imaginary part beta.

复变量 s 由实部α 和虚部β 组成。

[例 2] Most discrete-time signals *arise from* sampling continuous-time signals.

多数离散时间信号是由连续时间信号采样而来的。

② 与形容词连用，构成短语形容词，其语法功能类似过去分词，常用作后置定语，例如：analogous to, applicable to, capable of, different from, essential to, equivalent to, familiar with, full of, identical to, independent of, proportional to, similar to, suitable for, superior to 等。

[例 3] The power gain in decibels *is equal to* 10 times the logarithm base 10 of the power gain.

用分贝表示的功率增益等于 10 乘以以 10 为底的功率增益的对数。

[例 4] A CT signal is also called an analog signal because its waveform is often analogous to that of the physical variable.

由于连续时间信号的波形常与对应物理变量的波形相似，所以也常称其为模拟信号。

③ 与名词连用构成短语名词，例如：anything but, lack of, nothing but, numbers of, plenty of, reduction to, something wrong with, ratio … to …, advantage over, problem with 等。

[例 5] This section deals with the *advantages* of transistors *over* electron tubes.

本节讨论晶体管与电子管相比其所具有的优点。

2. 常用短语介词

短语介词可当做介词使用，常用短语介词包括 according to, along with, apart from, as to/for, because of, but for, by means of, due to, in accordance with, in addition to, in case of, in proportion to, in spite of, instead of, on account of, owing to, with reference to, with regard to, up to 等。

[例 6] The convolution integral and its use in fixed, linear system analysis *by means of* the principle of superposition are treated in Chapter 2.

第 2 章讲述应用叠加原理求解卷积积分及其在恒定线性系统分析中的应用。

[例 7] This book has all three major analysis methods: mesh, loop, and nodal, *along with* explanations of the advantages and disadvantages of each.

该书介绍了三种主要的电路分析方法，即网孔法、环路法与节点法，并比较其优劣。

3. 常用介词的特殊用法

① by

[例 8] *By* this definition, no work is done by holding a box in a hand. （根据，按照，后接表示规则、惯例、标准等词，常用在推导中）

根据这一定义，把一只盒子拿在手中并没有做功。

[例 9] Since power is proportional to the square of the voltage, if the input signal is increased *by* a factor of two, the output signal is increased *by* a factor of four. （与数词连用，表示相差的数量）

由于功率与电压的平方成正比，如果输入信号是原来的 2 倍，则输出信号为原来的 4 倍。

[例 10] *By* a family of curves *is meant* a specified set of curves which satisfy given conditions.

所谓曲线簇，是指能满足已知条件的一组特殊的曲线。

[例 11] In contrast, *multiplying* a signal *by* another signal is nonlinear. （与 multiply, divide 等连用表示“乘以，除以”）

相反，将一个信号与另一个信号相乘是非线性的。

② for

[例 12] *For the sake of* accuracy and convenience, the magnetic effect is utilized almost universally in electric measuring instrument. （常与 purpose, sake, interest, benefit, good 等连用，表示目的）

为准确与方便起见，电子测量仪器几乎普遍使用磁效应。

[例 13] It is necessary for us to solve this equation *for x*. （与 solve 连用表示方程所要求解的对象）

我们必须从这个方程解出 x。

③ in

[例 14] Semiconductor devices are very small *in* size and light *in* weight.（后接抽象名词，做状语）

半导体器材体积小，重量轻。

[例 15] Voltage is expressed numerically *in* volts.（以，用，后接表示方式的计量单位、语言等）

用伏特表示电压。

[例 16] The inclusion of Re causes a decrease *in* amplification.（在表示数量变化的名词后多用 in）

接入 Re 会使放大倍数降低。

④ of

[例 17] The study of sound *is of importance* not only in music and speech, but also in communication and industry. （表示描述关系，具有……（性质或特点），等效于对应的形容词，但语气更强）

研究声音不仅对音乐和语言，而且对通信和工业都有重要意义。

[例 18] The physicist made an important discovery in *the year of 1977*. （表示同位关系）

1977 年，这个物理学家有一个重要发现。

[例 19] A signal is a *description of* how one parameter varies with another parameter. （A of B 结构，A 和 B 为动宾关系，修饰的名词 A 由及物动词变来）

信号用以描述一个参量如何随另一个参量的变化而变化。

[例 20] The *change of* electrical energy into mechanical energy is done in motors.（A of B 结构，B 和 A 为主谓关系，修饰的名词 B 由不及物动词而来）

电能转换为机械能是通过电动机实现的。

⑤ on/upon

[例 21] The flip-flop will change the stored information only *upon* application of proper control signal. （一……就，在……之后，后接具有动词含义的名词或动名词）

只有施加了合适的控制信号后，该触发器才能改变存储的信息。

⑥ with

[例 22] Radio waves travel through most types of matter *with ease*. （等效于对应副词，语气更强）

无线电波通过大多数物质很容易。

[例 23] *With* a sine wave input, the technician would look to see if the output is also a sine wave.（对于，就……而言，用于句首）

技术员想要知道，对于正弦输入，输出信号是否也是正弦波。

[例 24] *With* all its disadvantages this design is considered to be one of the best.（尽管，与 all 连用）

尽管有其缺点，这一设计仍被认为是最佳设计之一。

[例 25] Lamps are lighted *with* switch on.（等效于 when the switch is turned on）

开关闭合灯点亮。

Exercises

1. Choose the best answer for each of the following questions.

（1）A continuous signal is also called ________ because its waveform is often analogous to that of the physical variable.

a. an analog signal

b. a discrete signal

c. a digital signal

d. a sampled signal

（2）A system is called linear if it has two mathematical properties: ____________.

a. continuous and shift-invariance

b. shift-invariance and homogeneity

c. additivity and homogeneity

d. additivity and shift-invariance

（3）The unit impulse signal is ________ the unit step signal .

a. the same as

b. the differential of

c. the integral of

d. the cumulative sum of

（4）Which of the following statements best describes the property of the system T if $y_1(t)=T[x_1(t)]$, $y_2(t)=T[x_2(t)]$ implies $T[ax_1(t-s)+bx_2(t-s)]=ay_1(t-s)+by_2(t-s)$, ________.

a. the system is linear

b. it is a linear time-invariant system

c. the system satisfies the principle of superposition

d. the system satisfies additivity and time-invariance

（5）Any signal can be represented as ______.

a. the sum of a series of sinusoids with the same amplitudes

b. the sum of a series of sinusoids at the same frequencies

c. the sum of a series of shifted and scaled unit step function

d. the sum of a series of shifted and scaled unit impulse function

(6) Which one of the following statements is not true? ______.

a. A system that multiplies the input signal by a constant is linear

b. A linear time-invariant system is completely characterized by the impulse response function $h(t)$

c. The electric circuit for squaring is an example for a linear time-invariant system

d. The principle of superposition is held for a linear system

2. Fill in the blanks. You may need to change some words slightly.

input	coefficient	function	component	decomposition
transform	sinusoid	response	frequency	impulse

The Fourier _____ is important because if we know the _____ of the system to sinusoids at many different _____, then we can use the same kind of trick we used with _____ to predict the response via the impulse response _____. First, we measure the system's response to _____ of all different frequencies. Next, we take out input (e.g., time-varying current) and use the Fourier _____ to compute the values of the Fourier _____. At this point the _____ has been broken down as the sum of its component sinusoids. Finally, we can predict the system's response simply by adding the responses for all the _____ sinusoids.

3. Translate into Chinese.

(1) If the input to a linear system is a sinusoidal wave, the output will also be a sinusoidal wave, and at exactly the same frequency as the input.

(2) Just as we can express any signal as the sum of a series of shifted and scaled impulses, so too we can express any signal as the sum of a series of (shifted and scaled) sinusoids at different frequencies.

(3) A signal is formally defined as a function of one or more variables that conveys information on the nature of a physical phenomenon.

(4) Nyquist sampling theorem states that if the highest-frequency component of a CT signal $x(t)$ is $f_{\max}$, then $x(t)$ can be recovered from its sampled sequence $x(nT)$ if the sampling frequency f_s is larger than $2f_{\max}$.

(5) In fact, if a discrete-time sequence is absolutely summable, then its spectrum is, as in the continuous-time case, a bounded and continuous function of ω.

(6) Systems are defined to be stable if every bounded input excites a bounded output. The condition for discrete-time systems to be stable is that all poles of $H(x)$ lie inside the unit circle on the z-plane.

4. Fill in blanks with the most appropriate choice.

(1) A system is defined ______ the relationship between two signals.

a. with　　b. by　　c. in terms of　　d. to

(2) The input signal _____ a system generates an output signal.

a. into　　b. to　　c. of　　d. through

（3）The input signal represents a physical process that is generated independently _____ the system.

a. of　　b. from　　c. in　　d. on

（4）Mathematical modeling is concerned _____ the development of different forms of equations that can be used to represent a system.

a. with　　b. in　　c. about　　d. of

（5）The question _____ which we address ourselves is to characterize the change in the signal properties using properties of the system.

a. in　　b. with　　c. to　　d. for

（6）This question is _____ fundamental importance in many physical processes of interest.

a. of　　b. at　　c. with　　d. in

（7）_____ the convolution theorem introduced in Chapter 2, the responses of linear, time invariant systems to arbitrary inputs can be computed if the unit-sample response is known.

a. In regard to　　b. According to　　c. Due to　　d. In addition to

（8）The magnitude of the output is the magnitude of the input multiplied _____ the magnitude of the frequency response.

a. with　　b. by　　c. to　　d. in terms of

（9）In other problems of signal and system analysis, our interest may be focused _____ designing systems to process signals in particular ways.

a. in　　b. with　　c. on　　d. to

（10）_____ enhancement and restoration, in many applications there is a need to design systems to extract specific pieces of information from signals.

a. Owing to　　b. In addition to　　c. With respect to　　d. Regardless of

Reading Material

System Identification — The Easy Case

Assume that someone brings you a signal processing system enclosed in a black box. The box has two connectors, one marked *input* and the other *output*. Other than these labels there are no identifying marks or documentation, and nothing else is known about what is hidden inside. What can you learn about such a system? Is there some set of measurements and calculations that will enable you to accurately predict the system's output when an arbitrary input is applied? This task is known as *system identification*.

You can consider system identification as a kind of game between yourself and an opponent. The game is played in the following manner. Your opponent brings you the black box (which may have been specifically fabricated for the purpose of the game). You are given a specified finite amount of time to experiment with the system. Next your opponent specifies a test input and ask you for your prediction — were this signal to be applied what output would result? The test input is now applied and your prediction put to the test.

This game has two levels of play. In this section we will learn how to play the easy version. The easy case is when you are given complete control over the black box.

Assume you are given one hour to examine the box in any way you wish (short of prying off the top). At the end of precisely one hour your opponent will reappear, present you with an input signal and ask you what you believe the box's response will be. The most straightforward way of proceeding would be to quickly apply as many different input signals as you can and to record the corresponding outputs. Then you win the game if your opponent's input signal turns out to be essentially one of the inputs you have checked. Unfortunately, there are very many possible inputs, and an hour is too short a time to test even a small fraction of them. To economize we can exploit the fact that the box contains a linear time-invariant system. If we have already tried input x_n there is no point in trying ax_n or x_{n-m}, but this still leaves a tremendous number of signals to check.

Our job can be made more manageable in two different ways, one of which relies on the time domain description of the input signal, and the other on its frequency domain representation. The frequency domain approach is based on Fourier's theorem that every signal can be written as the weighted sum (or integral) of basic sinusoids. Assume that you apply to the unknown system not every possible signal, but only every possible sinusoid. You store the system's response to each of these and wait for your opponent to appear. When presented with the test input you can simply break it down to its Fourier components, and exploit the filter's linearity to add the stored system responses with the appropriate Fourier coefficients.

Now this task of recording the system outputs is not as hard as it appears, since sinusoids are eigensignals of filters. When a sinusoid is input to a filter the output is a single sinusoid of the same frequency, only the amplitude and phase may be different. So you need only record these amplitudes and phases and use them to predict the system output for the test signal. For example, suppose the test signal turns out to be the sum of three sinusoids

$$x_n = X_1 \sin(\omega_1 n) + X_2 \sin(\omega_2 n) + X_3 \sin(\omega_3 n)$$

Then, since the filter is linear, the output is the sum of the three responses with the Fourier coefficients H_1, H_2 and H_3.

$$y_n = H_1 X_1 \sin(\omega_1 n + \phi_1) + H_2 X_2 \sin(\omega_2 n + \phi_2) + H_3 X_3 \sin(\omega_3 n + \phi_3)$$

More generally, any finite duration or periodic test digital signal can be broken down by the DFT into the sum of a denumerable number of complex exponentials

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \mathrm{e}^{i\frac{2\pi k}{N}n}$$

and the response of the system to each complex exponential is the same complex exponential multiplied by a number H_k, $H_k \mathrm{e}^{i\frac{2\pi k}{N}n}$. Using these H_k we can predict the response to the test signal

$$y_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k H_k \mathrm{e}^{i\frac{2\pi k}{N}n}$$

The H_k are in general complex (representing the gains and phase shifts) and are precisely the elements of the frequency response.

The above discussion proves that the frequency response provides a complete description of a filter. Given the entire frequency response (i.e., the response of the system to all sinusoids), we can always win the game of predicting the response for an arbitrary input.

The frequency response is obviously a frequency domain quantity; the duality of time and frequency domains leads us to believe that there should be a complete description in the time domain as well. There is, and we previously called it the *impulse response.*

Like the frequency response, the impulse response may be used to predict the output of a filter when an arbitrary input is applied. The strategy is similar to that we developed above, only this time we break down the test signal in the basis of SUIs rather than using the Fourier expansion. We need only record the system's response to each SUI, expand the input signal in SUIs, and exploit the linearity of the system. Unfortunately, the SUIs are not generally eigensignals of filters, and so the system's outputs will not be SUIs, and we need to record the entire output. However, unlike the frequency response where we needed to observe the system's output for an infinite number of basis functions, here we can capitalize on the fact that all SUIs are related by time shifts. Exploiting the time-invariance property of filters we realize that after measuring the response of an unknown system to a single SUI, we may immediately deduce its response to all SUIs! Hence we need only apply a single input and record a single response in order to be able to predict the output of a filter when an arbitrary input is applied! The set of signals we must test in order to be able to predict the output of the system to an arbitrary input has been reduced to a single signal! This is the strength of the impulse response.

The impulse response may be nonzero only over a finite interval of time but exactly zero for all times outside this interval. In this case we say the system has a *finite impulse response,* or more commonly we simply call it an FIR filter. The MA systems are FIR filters.

Let's explicitly calculate the impulse response for the most general causal moving average filter. Using the unit impulse as input yields

$$y_n = \sum_{l=0}^{L} g_l \delta_{n-L+l,0} = g_0 \delta_{n-L,0} + g_1 \delta_{n-L+1,0} + g_2 \delta_{n-L+2,0} + \ldots + g_{L-1} \delta_{n-1,0} + g_L \delta_{n,0}$$

which is nonzero only when $n = 0$ or $n = 1$ or $\cdots$ or $n = L$. Furthermore, when $n = 0$ the output is precisely $h_0 = g_L$, when $n = 1$ the output is precisely $h_1 = g_{L-1}$, etc., until $h_L = g_0$. Thus the impulse response of a general MA filter consists exactly of the coefficients that appear in the moving average sum, but in reverse order!

The impulse response is such an important attribute of a filter that it is conventional to reverse the definition of the moving average, and define the FIR filter via the *convolution* in which the indices run in opposite directions.

It is evident that were we to calculate the impulse response of the nonterminating convolution it would consist of the coefficients as well; but in this case the impulse response would never quite become zero. If we apply a unit impulse to a system and its output never dies down to zero, we say that the system is Infinite Impulse Response (IIR). IIR filters can indeed sustain an impulse response that is nonzero for an infinite amount of time. To see this consider the simple case

$$y_n = x_n + \frac{1}{2} y_{n-1}$$

For negative times n the output is zero, $y_n = 0$, but at time zero $y_0 = 1$, at time one $y_1 = \frac{1}{2}$ and thereafter y_n is halved every time. It is obvious that the output at time n is precisely $y_n = 2^{-n}$, which for large n is extremely small, but never zero.

Suppose we have been handed a black box and measure its impulse response. Although there may be many systems with this response to the unit impulse, there will be only one filter that matches, and the coefficients are precisely the impulse response in reverse order. This means that if we know that the box contains a filter, then measuring the impulse response is sufficient to uniquely define the system. In particular, we needn't measure the frequency response since it is mathematically derivable from the impulse response.

It is instructive to find this connection between the impulse response (the time domain description) and the frequency response (the frequency domain description) of a filter. The frequency response of the nonterminating convolution system

$$y_n = \sum_{i=-\infty}^{\infty} h_i x_{n-i}$$

is found by substituting a sinusoidal input for x_n, and for mathematical convenience we will use a complex sinusoid $x_n = \mathrm{e}^{\mathrm{j}\omega n}$. We thus obtain

$$H(\omega)x_n = y_n = \sum_{k=-\infty}^{\infty} h_k \mathrm{e}^{\mathrm{j}\omega(n-k)} = \sum_{k=-\infty}^{\infty} h_k \mathrm{e}^{-\mathrm{j}\omega k}\mathrm{e}^{\mathrm{j}\omega n} = H_k x_n$$

where we identified the Fourier transform of the impulse response h_k and the input signal. We have once again shown that when the convolution system has a sinusoidal input, its output is the same sinusoid multiplied by a (frequency-dependent) gain. This gain is the frequency response, but here we have found the FT of the impulse response; hence the frequency response and the impulse response are an FT pair. Just as the time and frequency domain representations of signals are connected by the Fourier transform, the simplest representations of filters in the time and frequency domains are related by the FT.

New Words

coefficient	系数
denumerable	可数的
duality	对偶性
duration	持续时间
eigensignal	特征信号
fabricate	制作，构成
nonterminating	无尽地
precisely	正好
capitalize on	利用
pry off	探查
system identification	系统辨识

FIR（Finite Impulse Response）	有限冲击响应
IIR（Infinite Impulse Response）	无限冲击响应
MA（moving average）	移动平均
SUI（sum of unit impulse）	单位冲击求和

Questions

1. What is called system identification?
2. How to predict the output of a filter if you are given the impulse response of it?
3. What is the difference between an FIR and an IIR filter?

Unit 5

Text

Radio Frequency and Microwave Applications

5.1 Introduction

This chapter lays the foundation for understanding higher-frequency wave phenomena and divides the task of active circuit design for RF/MW frequencies into specific concept blocks. These concept blocks create a gradual approach to understanding and designing RF/MW circuits and represent specific realms of knowledge that need to be mastered to become an accomplished designer.

Before we describe and analyze these types of waves we need to consider why RF/microwaves as a subject has become so important, that it is placed at the forefront of our modern technology[1]. Furthermore, we need to expand our minds to the many possibilities that these signals can provide for peaceful practices by exploring various commercial applications useful to mankind.

5.1.1 A Short History of RF and Microwaves

Circa 1864—1873, James Clark Maxwell integrated the entirety of man's knowledge of electricity and magnetism by introducing a set of four coherent and self-consistent equations that describe the behavior of electric and magnetic fields on a classical level. This was the beginning of microwave engineering, as presented in a treatise by Maxwell at that time. He predicted, purely from a mathematical standpoint and on a theoretical basis, the existence of electromagnetic wave propagation and that light was also a form of electromagnetic energy—both completely new concepts at the time.

From 1885 to 1887, Oliver Heaviside simplified Maxwell's work in his published papers. From 1887 to 1891, a German physics professor, Heinrich Hertz, verified Maxwell's predictions experimentally and demonstrated the propagation of electromagnetic waves. He also investigated wave propagation phenomena along transmission lines and antennas and developed several useful structures. He could be called the first microwave engineer.

Marconi tried to commercialize radio at a much lower frequency for long-distance communications, but as he had a business interest in all of his work and developments, this was not a purely scientific endeavor.

Neither Hertz nor Heaviside investigated the possibility of electromagnetic wave propagation inside a hollow metal tube because it was felt that two conductors were necessary for the transfer of electromagnetic waves or energy. In 1897, Lord Rayleigh showed mathematically that electromagnetic wave propagation was possible in waveguides, both circular and rectangular. He showed that there are infinite sets of modes of the TE and TM type possible, each with its own cut-off frequency. These were all theoretical predictions with no experimental verifications.

From 1897 to 1936, the waveguide was essentially forgotten until it was rediscovered by two men, George Southworth (AT&T) and W. L. Barron (MIT), who showed experimentally that a waveguide could be used as a small bandwidth transmission medium, capable of carrying high power signals.

With the invention of the transistor in the 1950s and the advent of microwave integrated circuits in the 1960s, the concept of a microwave system on a chip became a reality[2]. There have been many other developments, mostly in terms of application mass, that have made RF and microwaves an enormously useful and popular subject. Maxwell's equations lay the foundation and laws of the science of electromagnetics, of which the field of RF and microwaves is a small subset. *Due to the exact and all-encompassing nature of these laws in predicting electromagnetic phenomena, along with the great body of analytical and experimental investigations performed since then, we can consider the field of RF and microwave engineering a "mature discipline" at this time*[3].

5.1.2 Applications of Maxwell's Equations

As indicated earlier in Chapter 2, *Fundamental Concepts in Electrical and Electronics Engineering*, standard circuit theory can neither be used at RF nor particularly at microwave frequencies. This is because the dimensions of the device or components are comparable to the wavelength, which means that the phase of an electrical signal (e.g., a current or voltage) changes significantly over the physical length of the device or component. *Thus use of Maxwell's equations at these higher frequencies becomes imperative*[4].

In contrast, the signal wavelengths at lower frequencies are so much larger than the device or component dimensions, that there is negligible variation in phase across the dimensions of the circuit. Thus Maxwell's equations simplify into basic circuit theory, as covered in Chapter 3, *Mathematical Foundation for Understanding Circuits.*

At the other extreme of the frequency range lies the optical field, where the wavelength is much smaller than the device or circuit dimensions. In this case, Maxwell's equations simplify into a subject commonly referred to as geometrical optics, which treats light as a ray traveling on a straight line.

These optical techniques may be applied successfully to the analysis of very high microwave frequencies (e.g., high millimeter wave range), where they are referred to as "quasi-optical." Of course, it should be noted that further application of Maxwell's equations leads to an advanced field of optics called "physical optics or Fourier optics," which treats light as a wave and explains such phenomena as diffraction and interference, where geometrical optics fails completely.

The important conclusion to be drawn from this discussion is that Maxwell's equations present a unified theory of analysis for any system at any frequency, provided we use appropriate simplifications when the wavelengths involved are much larger, comparable to, or much smaller than the circuit dimensions[5].

5.1.3 Properties of RF and Microwaves

An important property of signals at RF, and particularly at higher microwave frequencies, is their great capacity to carry information. This is due to the large bandwidths available at these high frequencies. For example, a 10 percent bandwidth at 60 MHz carrier signal is 6 MHz, which is approximately one TV channel of information; on the other hand 10 percent of a microwave carrier signal at 60 GHz is 6 GHz, which is equivalent to 1000 TV channels.

Another property of microwaves is that they travel by line of sight, very much like the traveling of light rays, as described in the field of geometrical optics. Furthermore, unlike lower-frequency signals, microwave signals are not bent by ionosphere. Thus use of line-of-sight communication towers or links on the ground and orbiting satellites around the globe are a necessity for local or global communications.

A very important civilian as well as military instrument is radar. The concept of radar is based on radar cross-section which is the effective reflection area of the target. A target's visibility greatly depends on the target's electrical size, which is a function of the incident signal's wavelength. Microwave frequency is the ideal signal band for radar applications. Of course, another important advantage of use of microwaves in radars is the availability of higher antenna gains as the frequency is increased for a given physical antenna size. This is because the antenna gain being proportional to the electrical size of the antenna, becomes larger as frequency is increased in the microwave band. The key factor in all this is that microwave signal wavelengths in radars are comparable to the physical size of the transmitting antenna as well as the target.

There is a fourth and yet very important property of microwaves: the molecular, atomic, and nuclear resonance of conductive materials and substances when exposed to microwave fields. This property creates a wide variety of applications. For example, because almost all biological units are composed predominantly of water and water is a good conductor, microwave technology has tremendous importance in the fields of detection, diagnostics, and treatment of biological problems or medical investigations (e.g., diathermy, scanning, etc.). There are other areas in which this basic property would create a variety of applications such as remote sensing, heating (e.g., industrial purification and cooking) and many others that are listed in a later section.

5.2 Reasons for Using RF/Microwaves

Over the past several decades, there has been a growing trend toward use of RF/ microwaves in system applications. There are many reasons among which the following are prominent:

- Wider bandwidths due to higher frequency
- Smaller component size leading to smaller systems
- More available and less crowded frequency spectrum
- Better resolution for radars due to smaller wavelengths
- Lower interference due to lower signal crowding
- Higher speed of operation
- Higher antenna gain possible in a smaller space

On the other hand, there are some disadvantages to using RF/microwaves, such as: more expensive components, availability of lower power levels, existence of higher signal losses, and use of high-speed semiconductors (such as GaAs or InP) along with their corresponding less-mature technology (relative to the traditional silicon technology, which is now quite mature and less expensive).

In many RF/microwave applications the advantages of a system operating at these frequencies outweigh the disadvantages and propel engineers to a high-frequency design.

5.3 RF/Microwave Applications

The major applications of RF/microwave signals can be categorized as follows:

5.3.1 Communication

This application includes satellite, space, long-distance telephone, marine, cellular telephone, data, mobile phone, aircraft, vehicle, personal, and wireless local area network (WLAN), among others. Two important subcategories of applications need to be considered: TV and radio broadcast, and optical communications.

TV and radio broadcast

In this application, RF/microwaves are used as the carrier signal for audio and video signals. An example is the Direct Broadcast System (DBS), which is designed to link satellites directly to home users.

Optical communications

In this application, a microwave modulator is used in the transmitting side of a low-loss optical fiber with a microwave demodulator at the other end. The microwave signal acts as a modulating signal with the optical signal as the carrier. Optical communication is useful in cases where a much larger number of frequency channels and less interference from outside electromagnetic radiation are desired. Current applications include telephone cables, computer network links, low-noise transmission lines, and so on.

5.3.2 Radar

This application includes air defense, aircraft/ship guidance, smart weapons, police, weather, collision avoidance, and imaging.

5.3.3 Navigation

This application is used for orientation and guidance of aircraft, ships, and land vehicles. Particular applications in this area are as follows:

- Microwave Landing System (MLS), used to guide aircraft to land safely at airports
- Global Positioning System (GPS), used to find one's exact coordinates on the globe

5.3.4 Remote Sensing

In this application, many satellites are used to monitor the globe constantly for weather conditions, meteorology, ozone, soil moisture, agriculture, crop protection from frost, forests, snow thickness, icebergs, and other factors such as monitoring and exploration of natural resources.

5.3.5 Domestic and Industrial Applications

This application includes microwave ovens, microwave clothes dryers, fluid heating systems, moisture sensors, tank gauges, automatic door openers, automatic toll collection, highway traffic monitoring and control, chip defect detection, flow meters, power transmission in space, food preservation, pest control, and so on.

5.3.6 Medical Applications

This application includes cautery, selective heating, heart stimulation, hemorrhage control, sterilization, and imaging.

5.3.7 Surveillance

This application includes security systems, intruder detection, and Electronic Warfare (EW) receivers to monitor signal traffic.

5.3.8 Astronomy and Space Exploration

In this application, gigantic dish antennas are used to monitor, collect, and record incoming microwave signals from outer space, providing vital information about other planets, stars, meteors, and other objects and phenomena in this or other galaxies.

5.3.9 Wireless Applications

Short-distance communication inside as well as between buildings in a local area network(*LAN*) *arrangement can be accomplished using RF and microwaves*[6]. Connecting buildings via cables (e.g., coax or fiber optic) creates serious problems in congested metropolitan areas because the cable has to be run underground from the upper floors of one building to the upper floors of the other. This problem, however, can be greatly alleviated using RF and microwave transmitter/receiver systems that are mounted on rooftops or in office windows. Inside buildings, RF and microwaves can be used effectively to create a wireless LAN in order to connect telephones, computers, and various LANs to each other. Using wireless LANs has a major advantage in office rearrangement where phones, computers, and partitions are easily moved with no change in wiring in the wall outlets. This creates enormous flexibility and cost savings for any business entity.

Words & Expressions

active	['æktiv]	*adj.*	有源的
alleviate	[ə'li:vieit]	*vt.*	减轻；缓和
antenna	[æn'tenə]	*n.*	天线
carrier	['kæriə]	*n.*	载波
cautery	['kɔ:təri]	*n.*	烙（术）；烙器；烧灼剂
channel	['tʃænl]	*n.*	频道；信道
coax	[kəuks]	*n.*	同轴电缆
collision	[kə'liʒən]	*n.*	碰撞；抵触
coordinate	[kəu'ɔ:dinit]	*n.*	坐标（用复数）
diffraction	[di'frækʃən]	*n.*	衍射
entity	['entiti]	*n.*	实体；存在；本质
hemorrhage	['hemərid ʒ]	*n.*	出血
incident	['insidənt]	*adj.*	入射的，投在或射在表面上的
interference	[ˌintə'fiərəns]	*n.*	干扰
ionosphere	[ai'ɔnəsfiə]	*n.*	电离层

magnetism	['mægnitizəm]	*n.*	磁性；磁力
meteorology	[ˌmiːtjə'rɔlədʒi]	*n.*	气象学，气象状态
metropolitan	[metrə'pɔlɪt(ə)n]	*a.*	大都市的
optical	['ɔptikəl]	*adj.*	眼睛的；视力的；光学的
orientation	[ˌɔ(ː)rien'teiʃən]	*n.*	定位；定向
ozone	['əuzəun]	*n.*	新鲜的空气，[化]臭氧
propagation	[prɔpə'geiʃən]	*n.*	（声波，电磁辐射等）传播
sterilization	[sterilai'zeiʃən]	*n.*	杀菌，绝育
surveillance	[səː'veiləns]	*n.*	监视，监督
via	['vaiə, 'viːə]	*prep.*	经由；取道
cut-off frequency			截止频率
Transverse Magnetic (TM)			横磁（性）的
lay the foundation of			给……打下基础，为……奠定基础
line-of-sight			视线，瞄准线
modulating signal			调制信号

Notes

[1] Before we describe and analyze these types of waves we need to consider why RF/microwaves as a subject has become so important, that it is placed at the forefront of our modern technology. as a subject 做同位语，进一步说明 RF/microwaves.

在讲述这类电波之前，我们需要考虑这样一个问题：为什么射频/微波学科变得如此重要，以至于人们要将其归入现代技术的前沿学科？

[2] With the invention of the transistor in the 1950s and the advent of microwave integrated circuits in the 1960s, the concept of a microwave system on a chip became a reality. with 结构做状语，表示伴随状态。

随着 20 世纪 50 年代晶体管的发明和 60 年代微波集成电路的出现，芯片级微波系统的设想已经变成了事实。

[3] Due to the exact and all-encompassing nature of these laws in predicting electromagnetic phenomena, along with the great body of analytical and experimental investigations performed since then, we can consider the field of RF and microwave engineering a “mature discipline” at this time.

由于麦克斯韦定律可准确预测各种电磁现象，加之后续进行的大量理论分析和实验研究工作，可以说射频/微波工程现在是一门成熟的学科。

[4] Thus use of Maxwell’s equations at these higher frequencies becomes imperative. 利用转换法将名词 use 译为动词，并补充主语。

于是，人们迫切需要将麦克斯韦方程应用于这些更高的频段。

[5] The important conclusion to be drawn from this discussion is that Maxwell’s equations present a unified theory of analysis for any system at any frequency, provided we use appropriate simplifications when the wavelengths involved are much larger, comparable to, or much smaller than the circuit dimensions.

从以上讨论中可得如下重要结论：无论信号波长是远远大于或小于电路尺寸还是与之相近，只要进行适当的简化，麦克斯韦方程就可对任何频率下的任何系统提供统一的分析理论。

[6] Short-distance communication inside as well as between buildings in a local area network(LAN) arrangement can be accomplished using RF and microwaves. as well as 相当于 and。

利用射频和微波可实现局域网中同一建筑物内或不同建筑物间的短距离通信。

Grammar

As 的用法

As 在科技英语里应用非常活跃，可以用做关系代词、关系副词、连词和介词，还可以与其他词或词组构成固定搭配。

1. 用做介词，构成介词短语，在句中做同位语、状语或补足语

[例 1] Radio, *as the fastest and most reliable means of communication*, is essential to modern science. （同位语）

无线电是最为快捷可靠的通信工具，在现代科学中不可或缺。

[例 2] Electrical power is always carried over long distances *as* a high-tension voltage at low-current strength. （方式状语）

电力总是以高电压低电流的方式进行远距离输送。

[例 3] Maxwell's equations simplify into a subject commonly referred to *as* geometrical optics, which treats light *as* a ray traveling on a straight line. （补足语）

麦克斯韦方程可简化为通常所说的几何光学，它将光视为在直线上传播的光线。

注：某些及物动词或带介词的动词常常要求 as 引出补足语，在科技英语中常见的这类动词有：assume, define, describe, consider, regard, know, refer to 等。

2. 用作连词引导状语从句

[例 4] The antenna gain becomes larger *as* frequency is increased in the microwave band. （时间）

在微波波段，天线增益随着频率的升高而增大。

[例 5] Marconi tried to commercialize radio at a much lower frequency for long-distance communications, but *as* he had a business interest in all of his work and developments, this was not a purely scientific endeavor.（原因）

马可尼试图以更低的频率实现长距离通信，并使之商品化，但由于他所有的工作都涉及商业利益，所以这不是纯科学追求。

[例 6] Small *as* atoms are, electrons are still smaller. （让步）

原子虽然很小，但电子更小。

[例 7] *Just as* not all systems are linear, not all linear systems are shift-invariant. （方式）

正如系统并非都是线性系统一样，线性系统并非都是时不变系统。

[例 8] Electromagnetic waves travel *as* fast *as* light travels. （比较）

电磁波和光传播得一样快。

[例 9] The electrical resistance of a body is constant only *so long as* its physical condition is unchanged.

只要物体的物理条件不变，其电阻就是恒定的。（条件）

3. 用作关系代词或关系副词引导定语从句

[例 10] Compression, *as* the name implies, deals with techniques for reducing the storage required to save an image, or the bandwidth required to transmit it.

顾名思义，压缩研究减少图像存储容量或降低图像传输带宽的技术。

[例 11] The current is in *the same* direction *as* the motion of the positive particles. （引导限制性定语从句，还有 such… as…）

电流的方向即是正电荷运动的方向。

[例 12] Pulse-code modulation is *such as* the samples are quantized into discrete steps.

脉冲编码调制是抽样值被量化成离散的阶梯值的调制方法。

[例 13] *As* indicated earlier in Chapter 2, Fundamental concepts in electrical and electronics engineering, standard circuit theory can neither be used at RF nor particularly at microwave frequencies.

正如第 2 章“电气与电子工程中的基本概念”所指出，基本电路理论对射频，尤其微波频段不适用。

4. 固定短语结构

[例 14] QFSK effectively double the data rate *as against* binary that can be transmitted in a given bandwidth.

与二元制频移键控相比，正交频移键控在一定的带宽内使传输的数据速率实际上翻了一番。

[例 15] The key factor in all this is that microwave signal wavelengths in radars are comparable to the physical size of the transmitting antenna *as well as* target.

其中的关键因素是雷达中所用微波信号的波长与发射天线及目标尺寸相比拟。

[例 16] The major applications of RF/microwave signals are categorized *as follows*: communication, radar, navigation, remote sensing, domestic and industrial applications, and medical applications.

射频/微波的主要应用有：通信、雷达、导航、遥感、家庭与工业应用及医学应用等。

[例 17] *As far as* the principle of the conservation of energy *is concerned*, it is one of the general principles that underlie all natural process.

就能量守恒定理而言，它是构成一切自然过程的基础的普遍原理之一。

注：其他常见的固定短语结构还有 as a matter of fact, as a result (of), as a whole, as a rule, as for/to, as regards, so as to 等。

Exercises

1. Choose the best answer for each of the following questions.

（1）The beginning of microwave engineering is characterized by ________.

a. a piece of paper written by Oliver Heaviside

b. The discovery that electromagnetic wave propagation was possible in waveguide

c. The experimental demonstration of the propagation of electromagnetic waves

d. Maxwell's equations

（2）________ showed experimentally the possibility of electromagnetic wave propagation in waveguide?

a. Heinrich Hertz

b. George Southworth and W.L.Barron

c. Oliver Heaviside

d. Lord Rayleigh

（3）Basic circuit theory can be used at ____________.

a. FR/microwave frequency

b. the optical field

c. lower frequency

d. any frequency

（4）Which is not considered as the advantage to using RF/microwaves? ________.

a. larger bandwidth available

b. use of high-speed semiconductors along with their less-mature technology

c. higher antenna gain possible in a smaller space

d. better resolutions for radars due to smaller wavelengths

（5）RF/microwaves are used as _______ signal for audio signals in broadcasting.

a. modulated

b. carrier

c. modulating

d. demodulated

（6）Microwave clothes dryers are designed to take advantages of one of the following properties of microwaves? _______.

a. larger bandwidth available at high frequency

b. microwaves travel by line of sight

c. microwave signal wavelength are comparable to the physical size of clothes

d. resonance of conductive material when exposed to microwave field

2. Substitute the underlined words with the appropriate choices given below.

（1）Neither Herz nor Heaviside investigated the possibility of electromagnetic wave propagation inside a hollow metal tube.

a. transmission　b. reception　c. flow　d. transition

（2）An important property of signals at RF, and particularly at higher microwave frequency, is their great capacity to carry information.

a. volume　b. eligibility　c. capability　d. probability

（3）Almost all biological units are composed predominantly of water.

a. significantly　b. especially　c. fundamentally　d. primarily

（4）From 1897 to 1936 the waveguide was essentially forgotten until it was rediscovered by two men.

a. basically　　b. practically　　c. absolutely　　d. nearly

（5）In this application, gigantic dish antennas are used to monitor, collect and record incoming microwave signals from outer space.

a. observe　　b. supervise　　c. measure　　d. process

（6）In many RF/microwave applications, the advantages of a system operating at these frequencies outweigh the disadvantages and propel engineers to a high-frequency design.

a. exceed　　b. extend　　c. expand　　d. express

3. Translate the following sentences into Chinese.

（1）Microwaves are normally considered to embrace the frequency range 109-1012 Hz or a characteristic wavelength range of 30 cm to 0.3 mm.

（2）The study of electromagnetic radiation is an exact science because it can be represented exactly by mathematical expressions.

（3）Microwaves are necessary for communication with satellites because they can pass through the ionosphere which reflects lower frequency radio waves.

（4）The rate of microwave power absorption in most materials is proportional to its water content. This property can be used to provide microwave heating. Because the microwave signal penetrates most non-conductors, microwave power provides a most efficient means of applying heat uniformly throughout a body.

（5）Microwaves are potentially hazardous because of their heating effect. The effect may not be felt until damage has already been done because the heating may be internal whereas our body is designed to warn us about externally applied heat.

（6）For seismic signals traveling through a portion of the earth and picked up by appropriate transducers, one can extract from their particular shape and appearance information concerning the underlying strata.

4. Complete the sentences using "as structure".

（1）We consider silver ________.

我们认为银是良导体。

（2）Another advantage of use of microwaves in radars is the availability of higher antenna gain ________ for a given physical antenna size.

在雷达上使用微波的另一个优点是，对于给定的天线尺寸，随着频率的升高可获得更大的天线增益。

（3）In this application, RF/microwaves are used ________.

在此应用中，射频/微波用作音频和视频信号的载波信号。

（4）________, their rate is gradually halved.

当这些脉冲经过微芯片的电子电路时，它们的速率逐步减半。

（5）Short-distance communication inside ________ in a LAN arrangement can be accomplished using RF and microwaves.

局域网中建筑物内及其间的短距离通信可以用射频和微波实现。

（6）With the development of computers it is ________of the mind.

随着计算机的问世，人似乎突然成了智能上的百万富翁。

（7）The surface of the earth has not always looked ________ today, in fact the earth is changing everyday.

地球表面的情况并不总是一成不变的，实际上地球每日都在变化着。

（8）Without sound, our world would be ________.

没有声音，世界就会变得如同无声电影一样奇怪和不自然。

（9）________, microwave radars are used for this purpose.

由于普通雷达不能探测很小的物体，微波雷达就专供此用。

（10）Another property of microwaves is that they travel by line of sight, very much like ________.

微波的另一特性是它们是视线传播，很像光线的传播，正如在几何光学里所描述的一样。

Reading Material

Basic RF Building Blocks

RF systems are constructed primarily using four basic building blocks — amplifiers, filters, mixers, and oscillators. Amplifiers and filters are common analog blocks and are well handled by SPICE. However, mixers and oscillators are not heavily used in analog circuits and SPICE has limited ability to analyze them. What makes these blocks unique is presented next.

1. Mixers

Mixers translate signals from one frequency range to another. They have two inputs and one output. One input is for the information signal and the other is for the clock signal, the LO. Ideally, the signal at the output is the same as that at the information signal input, except shifted in frequency by an amount equal to the frequency of the LO. A multiplier can act as a mixer. In fact, a multiplier is a reasonable model for a mixer except that the LO is passed through a limiter, which is usually an integral part of the mixer, to make the output less sensitive to noise on the LO.

The input and output signals of a mixer used for up-conversion (as in a transmitter) are shown in Figure 1. The LO is shown after passing through the limiter so that the output in the time-domain is simply the product of the inputs, or the convolution of the two inputs in the frequency domain. The information signal, here a modulation signal, is replicated at the output above and below each harmonic of the LO. These bands of signal above and below each harmonic are referred to as sidebands. There are two sidebands associated with each harmonic of the LO. The ones immediately above the harmonics are referred to as the upper sidebands and the ones below are referred to as the lower sidebands. The sideband at DC is referred to as the baseband.

When the LO has a rich harmonic content, an input signal at any sideband will be replicated to each of the sidebands at the output. Usually, only one sideband is of interest and the others must be eliminated. If the desired sideband is the baseband, then the undesired sidebands are eliminated with a lowpass filter. Otherwise the undesired sidebands are removed with a bandpass filter. This works well for sidebands of harmonics different from that of the desired sideband. However, special techniques are then required to eliminate the remaining undesired sideband.

Consider a down-conversion mixer (as in a receiver) and assume the mixer is followed by a filter. This filter is used to remove all but the desired channel. The output of the mixer/filter pair is sensitive to signals in each sideband of the LO. Associated with each sideband is a transfer function from that sideband to the output. The shape of the transfer function is determined largely by the filter. Thus, the bandwidth of the passband is that of the filter. If the filter is a bandpass, then the passband of the transfer function will be offset from the LO or its harmonic by the center frequency of the filter. These passbands are referred to as the images of the filter and are shown in Figure 2. Generally only one image is desired, the rest are undesired. The most troubling is usually the one that shares the same harmonic as the desired image. Image-reject mixers are designed to reduce the gain associated with this undesired image.

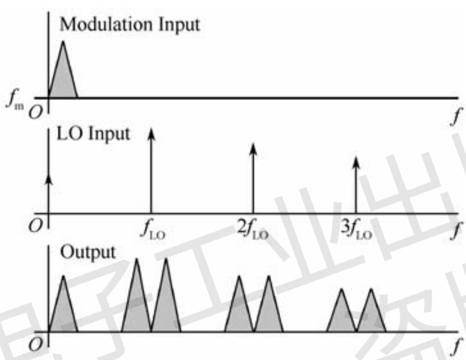


Figure 1 Signals at the inputs and outputs of an up-conversion mixer. The modulation signal is mixed up to the upper and lower sidebands of the LO and its harmonics.

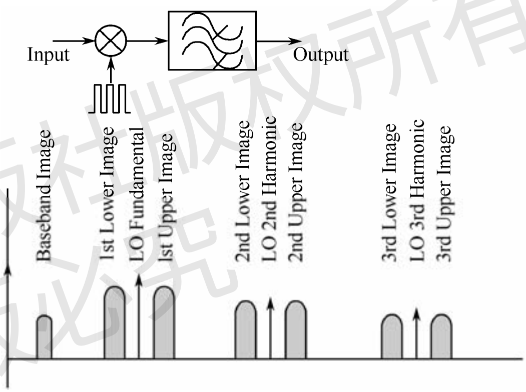


Figure 2 Images at the input of the first mixing stage of a typical receiver. The images are frequency bands where the output is sensitive to signals at the input.

Sidebands and images are related, but are not the same. Sidebands are frequency bands in the signal actually produced at the output of a mixer, whereas images are bands at the input of a mixer that have the potential to produce a response at the output frequency.

2. Oscillators

Oscillators generate a reference signal at a particular frequency. In some oscillators, referred to as VCOs for voltage controlled oscillators, the frequency of the output varies proportionally to some input signal. Compared to mixers, oscillators seem quite simple. That is an illusion.

Oscillators are generally used in RF circuits to generate the LO signal for mixers. The noise performance of the mixer is strongly affected by noise on the LO signal. The LO is always passed through a limiter, which is generally built into the mixer, to make the mixer less sensitive to small variations in the amplitude of the LO. However, the mixer is still sensitive to variations in the phase of the LO. Thus, it is important to minimize the phase noise produced by the oscillator.

Nonlinear oscillators naturally produce high levels of phase noise. To see why, consider the trajectory of an oscillator's stable periodic orbit in state space. Furthermore, consider disturbing the oscillator by applying an impulse $\mu(t)=\delta(t)$ .The oscillator responds by following a perturbed trajectory $v(t)+\Delta v(t)$ as shown in Figure 3, where $v(t)$ represents the unperturbed T-periodic solution and $\Delta v(t)$ is the perturbation in the response.

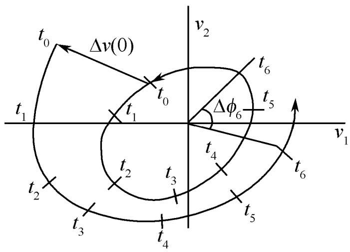


Figure 3 The trajectory of an oscillator shown in state space with and without a perturbation Δv . By observing the time stamps $(t_0 \cdots t_6)$ one can see that the deviation in amplitude dissipates while the deviation in phase does not.

Decompose the perturbed response into amplitude and phase variations

$$v_n(t)=v(t)+\Delta v(t)=(1+\alpha(t))v\left(t+\frac{\phi(t)}{2\pi f_0}\right) \tag{1}$$

where $v_n(t)$ represents the noisy output voltage of the oscillator, $\alpha(t)$ represents the variation in amplitude, $\phi(t)$ is the variation in phase, and $f_0={}^{1}\!/\!_{T}$ is the oscillation frequency.

Since the oscillator is stable and the duration of the disturbance is finite, the deviation in amplitude eventually decays away and the oscillator returns to its stable orbit. In effect, there is a restoring force that tends to act against amplitude noise. This restoring force is a natural consequence of the nonlinear nature of the oscillator and at least partially suppresses amplitude variations, as shown in Figure 4. With linear oscillators, there is no restoring force and so the amplitude is arbitrary (i.e., they do not have stable orbits). As such, linear oscillators exhibit equal amounts of amplitude and phase noise because the amplitude noise is not suppressed.

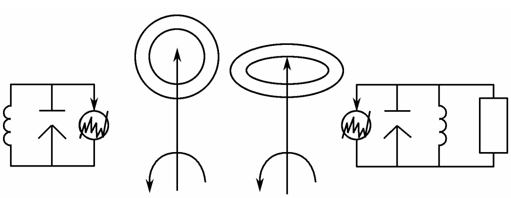


Figure 4 A linear oscillator along with its response to noise (left) and a nonlinear oscillator with its response to noise (right). The arrows are phasors that represents the unperturbed oscillator output, the carriers, and the circles represent the response to perturbations in the form of noise. With a linear oscillator the noise simply adds to the carrier. In a nonlinear oscillator, the nonlinearities act to control the amplitude of the oscillator and so to suppress variations in amplitude, thereby radially compressing the noise ball and converting it into predominantly a variation in phase.

Since the oscillator is autonomous, any time-shifted version of the solution is also a solution. Once the phase has shifted due to a perturbation, the oscillator continues on as if never disturbed except for the shift in the phase of the oscillation. There is no restoring force on the phase and so phase deviations accumulate. This is true for both linear and nonlinear oscillators. Notice that there is only one degree of freedom — the phase of the oscillator as a whole. There is no restoring force when the phase of all signals associated with the oscillator shift together; however there would be a restoring force if the phase of signals shifted relative to each other. This is important in oscillators with multiple outputs, such as quadrature oscillators or ring oscillators. The dominant phase variations appear identically in all outputs, whereas relative phase variations between the outputs are naturally suppressed by the oscillator, or added by subsequent circuitry and so tend to be much smaller.

New Words

amplitude	振幅
autonomous	自发的
deviation	背离
harmonic	谐波，谐函数
image	镜像
limiter	限幅器
mixer	混频器
modulation	调制
oscillator	振荡器
passband	通频带
phasor	相量
spectrum	光谱，频谱
trajectory	轨迹
bandpass filter	带通滤波器
LO（Local oscillator）	本地振荡器
VCO（voltage controlled oscillator）	压控振荡器

Questions

1．Describe both sidebands and images phenomena appeared in the mixer.
2．List some of the use of oscillators.
3．Why do nonlinear oscillators naturally produce higher levels of phase noise?