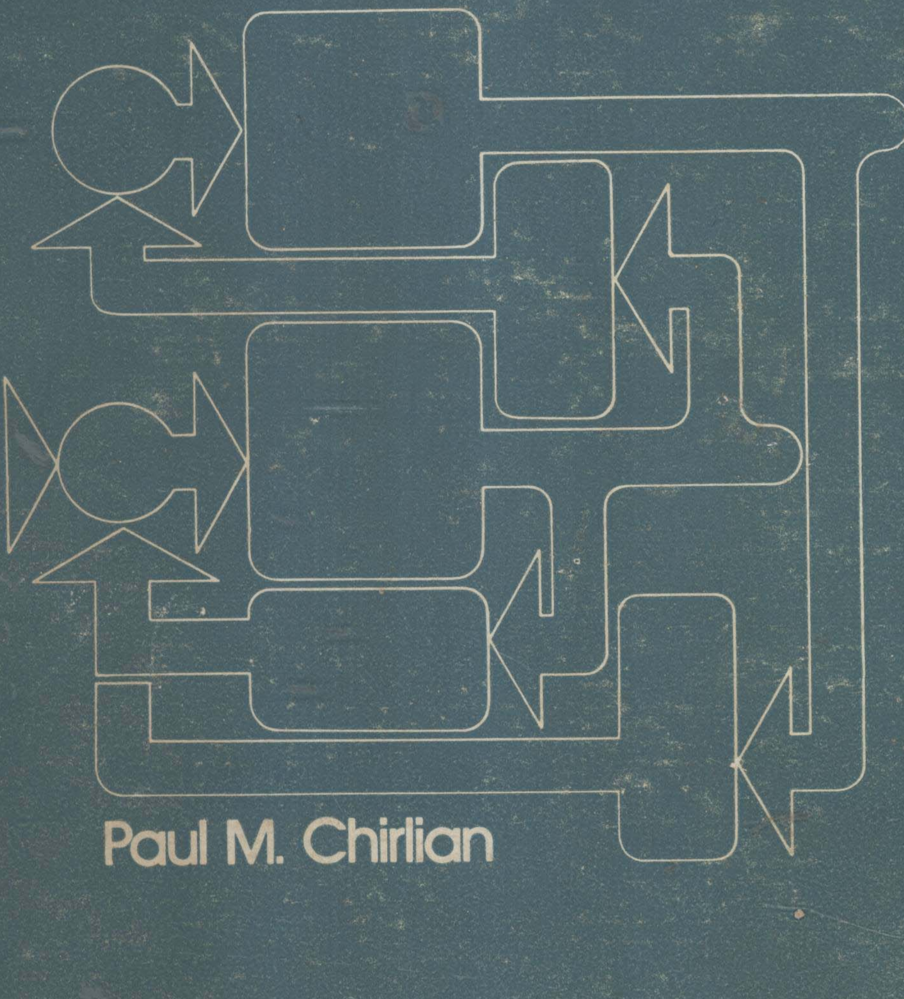
Signals, Systems, and the Computer



Paul M. Chirlian

SIGNALS, SYSTEMS, AND THE COMPUTER

PAUL M. CHIRLIAN
Stevens Institute of Technology

INTEXT EDUCATIONAL PUBLISHERS
New York and London

To Barbara, Lisa, and Peter

The Intext Series in

Circuits, Systems, Communications, and Computers

Consulting Editor

S. C. Gupta

Chirlian—SIGNALS, SYSTEMS, AND THE COMPUTER

Gupta, Bayless, and Peikari—CIRCUIT ANALYSIS: With Computer Application to Problem Solving

Ingels—INFORMATION AND CODING THEORY

Lathi—SIGNALS, SYSTEMS, AND CONTROLS

Matsch—ELECTROMAGNETIC AND ELECTROMECHANICAL MACHINES

Mickle and Sze—OPTIMIZATION IN SYSTEMS ENGINEERING

Sheng—INTRODUCTION TO SWITCHING LOGIC

Library of Congress Cataloging in Publication Data

Chirlian, Paul M

Signals, systems, and the computer.
(Intext series in circuits, systems, communication, and computers)
Includes bibliographies.
1. Electronic data processing—Electric networks.
2. Electronic data processing—System analysis.

I. Title.

TK454.2.C46 621.3819'535 72-84133

ISBN 0-7002-2427-0

Intext Educational Publishers
257 Park Avenue South
New York, New York 10010

Preface

This textbook on signals and systems is intended for use in senior or first-year graduate electrical engineering courses. It is designed to follow a basic network theory course and to provide the student with the mathematical tools needed to analyze a wide variety of systems and the signals processed by them. It may be used for a one or two semester course since material may be considered or omitted according to the instructor's requirements.

Much of the difficulty of such analyses has been greatly reduced by the application of the digital computer to the solution of the mathematical problems. However, the digital computer does not eliminate the need for a thorough understanding of the principles of analysis which are the primary concern of the book. Nevertheless, the student should feel confident that he will be able to write a computer program using FORTRAN IV that can implement the solution of a mathematical problem.

This is not intended to be a handbook of computer programs, hence a computer program is not given for every analysis encountered. Moreover, since this is not a study of numerical analysis, not every common type of numerical analysis is included. Methods for minimizing storage space are discussed and flow charts of many of the programs are included. Programs written for batch processing and timeshared operation are also covered because of the rapidly increasing use of these techniques. Thus we shall at times write programs which are suitable for these operations.

In Chapter 1 the general concepts of systems, and the signals transmitted by them are discussed. The basic idea of digital computation is briefly presented. A direct-current loop-analysis program is written and explained to illustrate how a FORTRAN IV program can simplify one of the most tedious problems that the student has encountered.

In Chapter 2 the Fourier transform is presented using distribution theory. An elementary discussion of distribution theory is covered, which can be easily understood by a student who is unfamiliar with the subject. The theory greatly simplifies the Fourier transform of certain common functions and allows generalized functions, such as the unit impulse and its derivatives, to be treated rigorously but simply. The Fourier transform is derived and its fundamental properties are discussed, and the Fourier series is treated as a special case of the Fourier transform. In addition, the fast Fourier transform and a computer program for implementing it are discussed in great detail.

The convolution theorem, Gibbs phenomenon, and Shannon's sampling theorem are also included.

In Chapter 3 the Laplace transform is derived from the Fourier transform and the ideas of distribution theory are applied. The 0−, rather than the 0+, Laplace transform is used and the advantages of this transform in computing the response to an impulse applied at $t = 0$ are thoroughly discussed. The basic theorems and concepts of the Laplace transform used in the solution of differential equations are considered. The inverse Laplace transform is developed, but this may be omitted. It is assumed that the reader is familiar with the basic notions of functions of a complex variable. (Appendix B presents this material.) The convolution theorem and the two-sided Laplace transform are also discussed.

In Chapter 4 the basic principles of state space are considered and the procedures for writing state variable equations are given. Linear, time-invariant, time-varying, and nonlinear systems are discussed and techniques for the solution of state-variable equations are given. Numerical techniques are also discussed in detail and FORTRAN IV programs, which implement these solutions, are presented.

In Chapter 5 the analysis techniques of Chapters 2, 3, and 4 are applied to linear continuous time systems. The relation among transfer functions, impulse response, and sinusoidal steady-state response is given. The response to arbitrary signals in terms of unit step and unit impulse response is discussed and causal sytems are considered. The Hilbert transforms and Bode relations are derived as a consequence of causal systems. Low pass and band pass systems are also discussed. In addition, bounds on system response and the meaning of the effective bandwidth of a signal are considered.

In Chapter 6 discrete time systems are considered. Difference equations are discussed and state variable procedures are presented. Linear, time-invariant, time-varying, and nonlinear systems are considered. The solution of these equations, including computer implementation, is presented. The use of the *z*-transform for the solution of linear, time-invariant systems is discussed.

In Chapter 7 system stability is covered. Linear, time-invariant systems are considered first. The Routh-Hurwitz algorithms and Nyquist criterion are derived. Then, general (nonlinear) systems are considered. State variable procedures are discussed. Liapunov stability is presented in detail. Observability and controllability are discussed. Stability in sampled systems is also considered.

In Chapter 8 basic ideas of probability are presented. Random signals and processes are discussed. A simplified derivation of the central limit theorem is given. The basic ideas of spectral density are discussed. The effects of band pass filtering upon noise probability are covered. Correlation functions and the Weiner-Kinchine relation are discussed. A computer program for the implementation of correlation and the use of correlation to extract signals from noise are discussed.

In Chapter 9 transmission of information is covered and the basic ideas of information theory are used to develop the concepts of channel capacity.

Some basic ideas of encoding are also discussed. The transmission of signals over noisy channels and the extraction of the signals from the noise is considered. Noise-reducing codes are discussed. Next the use of continuous filtering to extract signals from noise is presented, and predicting and causal filters are discussed.

In Chapter 10 distributed systems are presented, and partial differential equations of distributed systems are derived and solved using Laplace transforms. The transient response of special and general transmission lines is discussed. Sinusoidal, steady-state response and standing waves are then considered. A discussion of impedance calculations using the Smith chart is presented. Stub matching is discussed and a computer program to implement this is obtained.

There are three appendixes for review or instruction of material which may be unfamiliar to the reader. Othogonal functions are covered in Appendix A, basic complex variable theory is discussed in Appendix B, and matrices are considered in Appendix C, which includes the Cayley-Hamilton theorem.

My loving and heartfelt thanks, and great appreciation, are given to my wife, Barbara, for typing and correcting the numerous drafts of the manuscript for this book. The author also wishes to thank his colleagues Professors A. C. Gilmore, Jr., G. J. Herskowitz, E. Peskin, H. W. Phair, and S. Smith, and Professors S. C. Gupta and L. Gerhardt for their invaluable suggestions.

Contents

CHAPTER 1

Introduction to Signals and Systems—The Digital Computer in Signal and System Analysis

In this book we shall study the analysis of *systems* and *signals*. We shall define a system as *a collection of devices which perform some specified objective*. Systems can be very simple or extremely complex. A flashlight consisting of a battery, switch, light, reflector, and case can be considered to be a system. A satellite communications link, which includes transmitters, receivers, satellite, computers, and antennas, is also a system. An electric network can also be considered a system. Thus, the complexity of systems may vary greatly.

Systems of the type that we shall consider respond to certain inputs. These inputs and the system's response, or outputs, will be called *signals*. In the communication system discussed, the input signal could be a signal from a microphone and the output signal the sound from a loud-speaker. A complex system can be assumed to be made up of components each of which is often treated as a system. There are many definitions of signals and systems. For this reason, mathematical procedures are usually used in classifying them.

The digital computer is often used in systems analyses to avoid tedious calculations. Accordingly, we shall often discuss computer programs which can implement analysis procedures. While certain analysis techniques lend themselves to computer solutions and others do not, all of these techniques are important since they usually provide additional insight into the operation of a system.

1-1. SOME FUNDAMENTAL ASPECTS OF SYSTEMS

Many systems can be represented by a set of equations which relate the input and output signals. Often, these will be differential equations. For instance, consider the system characterized by the "black box" of Fig. 1-1. There are k inputs $y_1, y_2, y_3, \ldots, y_k$ and n outputs $x_1, x_2, x_3, \ldots, x_n$. All of these are functions of time. For example, in an electrical network the $y(k)$ could be voltage generators and the $x(k)$ loop currents. If the system is linear, a set of differential equations which characterize the system could be

$$a_{1m}\frac{d^m x_1(t)}{dt^m} + \cdots + a_{11}\frac{dx_1(t)}{dt} + a_{10}x_1(t) = F_1[y_1(t), y_2(t), \ldots, y_k(t)]$$

$$a_{2m}\frac{d^m x_2(t)}{dt^m} + \cdots + a_{21}\frac{dx_2(t)}{dt} + a_{20}x_2(t) = F_2[y_1(t), y_2(t), \ldots, y_k(t)] \quad (1\text{–}1)$$

$$\text{- - - - - - - - - - - - - - - - - -}$$

$$a_{nm}\frac{d^m x_n(t)}{dt^m} + \cdots + a_{n1}\frac{dx_n(t)}{dt} + a_{n0}x_n(t) = F_n[y_1(t), y_2(t), \ldots, y_k(t)]$$

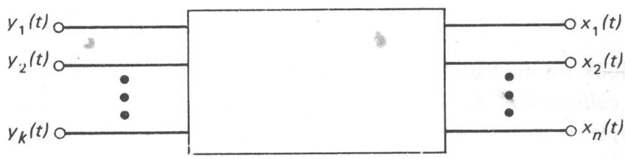


FIG. 1–1
A "black box" representation of a system with k inputs and n outputs

The a_{ij} are coefficients which may or may not be functions of time.

Let us now consider some terminology:

The *order of a system* is the highest order of the derivative necessary to characterize the system.

The *degree of a system* is the number of simultaneous equations needed to characterize the system.

Linear and Nonlinear Systems. Let us discuss a system with k inputs $y_1(t)$, $y_2(t), \ldots, y_k(t)$ and n outputs $x_1(t), x_2(t), \ldots, x_n(t)$. Suppose a particular set of inputs $y_{1a}(t), y_{2a}(t), \ldots, y_{ka}(t)$ is applied and that the response to them is $x_{1a}(t), x_{2a}(t), \ldots,$ $x_{na}(t)$. If another set of inputs $y_{1b}(t), y_{2b}(t), \ldots, y_{kb}(t)$ is applied, then the response to them is $x_{1b}(t), x_{2b}(t), \ldots, x_{nb}(t)$. Now suppose that the input signal becomes $y_{1a}(t) + by_{1b}(t), y_{2a}(t) + by_{2b}(t), \ldots, y_{ka}(t) + by_{kb}(t)$, where b is an arbitrary constant; i.e., each input signal becomes the sum of the original input plus b times the second one. The system is a *linear* one if the output is

$$\begin{aligned} x_1(t) &= x_{1a}(t) + bx_{1b}(t) \\ x_2(t) &= x_{2a}(t) + bx_{2b}(t) \\ &\text{- - - - - - - - - -} \\ x_n(t) &= x_{na}(t) + bx_{nb}(t) \end{aligned} \quad (1\text{–}2)$$

for all y's and b's. One consequence of Eqs. 1–2 is that a linear system is one whose output, due to a sum of (sets of) inputs, is the sum of the outputs which result when each (set of) input(s) acts separately. Such a system is said to satisfy the principle of *superposition*. For instance, in a linear electrical resistance if $i_1(t)$ results from the application of $v_1(t)$, and $i_2(t)$ results from the application of $v_2(t)$, then $i_1(t) + i_2(t)$ will result from the application of $v_1(t) + v_2(t)$.

Another consequence of Eqs. 1–2 is that if all inputs are multiplied by a constant, then all the outputs will be multiplied by the same constant. Such a system is said to be *homogenous.* In the case of the linear resistance, if $i_1(t)$ results from the application of $v_1(t)$, then $5i_1(t)$ will result from the application of $5v_1(t)$. If a system is linear, then it is characterized by linear simultaneous differential equations.

If a system is not linear, it is said to be *nonlinear.* A nonlinear system is characterized by a set of nonlinear differential equations. That is, there are products of variables, etc. A resistance whose voltage is given by $5i^3(t)$ is nonlinear. Also transistors are nonlinear devices.

Time-Invariant and Time-Varying Systems. A time-invariant system is one whose parameters do not change with time. As an example, consider Fig. 1–2. If the switch remains open (or closed) for all values of time, then this network is time invariant. However, if the switch is open for one value of time and then closed for another (e.g., open for $0 \leqq t \leqq 1$, closed for $1 < t \leqq 2$, open for $2 < t \leqq 5$, etc.), then the network is said to be time varying.

Note that the signals in this case $v_1(t)$ and $v_2(t)$, will be functions of time even if the switch remains fixed. However, this does not influence whether or not the system is time varying or time invariant.

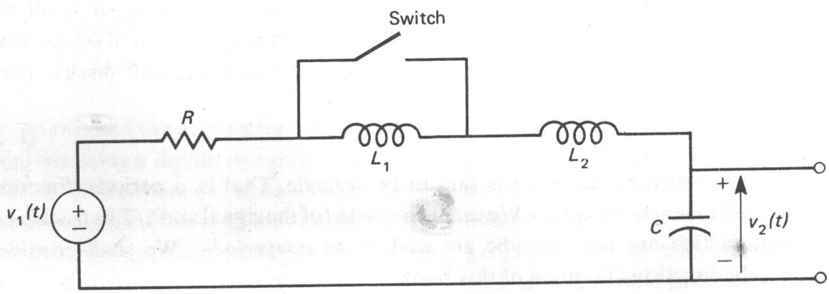


FIG. 1–2
A simple system

Continuous Time and Discrete Time Systems. Systems characterized by a set of differential equations, such as those of Eqs. 1–1 are called *continuous time systems.* That is, the inputs and outputs consist of functions of time which can vary at all times. The outputs consist of similar functions of time. In general, all systems are of this type. However, there are systems whose inputs and outputs are *satisfactorily approximated* if they are measured (or determined) only at discrete times, i.e., every second or every 10 seconds, etc. These are called *discrete time systems.* There are special techniques which simplify the analysis of these systems. We shall discuss them in Chapters 6 and 7, where these systems will be represented by *difference equations.*

Instantaneous Systems and Dynamic Systems. The response of most systems depends upon all its past history. Consider the network of Fig. 1–2. The value of $v_2(t)$ at $t = t_0$ is a function of $v_1(t)$ for all $t \leqq t_0$. This is called a *dynamic system.* Now consider an electric network made up only of linear resistors and voltage generators. At any instant of time t_0, the voltage and currents only depend upon the generator voltage at t_0; e.g., $i(t_0) = v(t_0)/R$. This is called an *instantaneous system.* In general, most practical systems are dynamic, and we shall mostly study this type of system.

There are other classifications of systems. For instance, the network of Fig. 1–2 is made up of lumped elements (resistors, inductors, and capacitors). This can be called a *lumped-parameter system.* If the elements of a system are distributed continuous, as in a transmission line, then we speak of it as a *distributed-parameter system.*

1-2. SOME FUNDAMENTAL ASPECTS OF SIGNALS

The inputs to a system and its responses to them are called *signals.* Many of the signals that we shall consider will be electrical. The input to a high-fidelity amplifier system can be considered to be the electrical output of a phonograph pickup. Similarly, if the system is the suspension of an automobile, the input is all the mechanical forces applied to the wheels, while the output is the motion of the car seat.

Signals can be classified in many ways, some of which we shall discuss here. If a signal $f(t)$ is such that

$$f(t) = f(t + T), \quad \text{for all} \quad t \tag{1–3}$$

where T is a constant, then $f(t)$ is said to be *periodic.* That is, a periodic function repeats itself for each T seconds. We call T the *period* of the signal and $1/T$ its *frequency.* Functions which are not periodic are said to be *nonperiodic.* We shall consider nonperiodic functions in much of this book.

Some signals are known or can be predicted for all time. For instance, suppose the input to a system is

$$\begin{aligned} f(t) &= e^{-t}, & t > 0 \\ f(t) &= 0, & t < 0 \\ &= \tfrac{1}{2}, & t = 0 \end{aligned} \tag{1–4}$$

Then $f(t)$ is known for all times, future, and past.

At times we deal with signals whose future values are unknown and cannot be exactly predicted. For instance, a noise signal which results from the random motion of charge carriers in a semiconductor is such a signal. Signals of this type are called *random.* If $f(t)$ is known for all future times, then it is said to be *nonrandom, deterministic,* or *predictable.* Of course, all real signals are random to some extent. For instance, we cannot predict for all future time what any real signal will be. However, it often is convenient to hypothesize mathematically that a predictable signal is applied to a system and then to compute its response. This often provides much information about the general system response.

Sometimes we must work with random signals whose values as functions of time are not known. However, the total energy contained in the signals can often be specified and probabilistic information is known about the signals. (This is discussed in Chapters 8 and 9.)

1-3. COMPUTER SOLUTION OF SIGNAL AND SYSTEM PROBLEMS

The mathematical analysis of system problems is often tedious and time consuming. A digital computer can be utilized to reduce greatly the calculation time and work. Hence, we shall, where practical, discuss computer programs that can be used to implement the analysis procedures considered in this book. These programs will be written in FORTRAN IV. Programs written for batch processing and timeshared operation are very similar. The only differences are minor differences in the input and output statements. Because of its rapidly increasing use, we shall at times write programs which are suitable for timeshared operation.

1-4. COMPUTER SOLUTION OF SIMULTANEOUS EQUATIONS

In this and the next section we shall discuss the procedure for performing dc loop analysis using a digital computer. Here we shall consider a program for the solution of a set of simultaneous equations, and in the next section we shall discuss computer procedures for obtaining the simultaneous equations from the network.

The mathematical basis of any computer program must be understood before the program is considered. We shall first discuss a mathematical procedure for the solution of a set of simultaneous equations. The computer program will be considered subsequently. We shall use the *Gauss-Jordan* method for the solution of a set of simultaneous equations. Consider the following equations:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\
&\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= y_n
\end{aligned} \tag{1-5a}$$

This can be written in matrix form as

$$\hat{a}\hat{x} = \hat{y} \tag{1-5b}$$

Note that the "hat" indicates a matrix. In some texts boldface letters are used to indicate a matrix. This is not done here since boldface letters are used to denote complex quantities. (A brief discussion of matrices is given in Appendix C.) The a_{ij} and the y_j are knowns and we must determine the x_j. Multiply the first equation by $-a_{21}/a_{11}$ and add the result to the second equation. Then repeat this for each equation in turn; e.g., multiply the first equation by $-a_{31}/a_{11}$ and add the result to the third one. The resulting set of equations becomes

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1$$

$$0 + \left(a_{22} - \frac{a_{12}a_{21}}{a_{11}}\right)x_2 + \cdots + \left(a_{2n} - \frac{a_{1n}a_{21}}{a_{11}}\right)x_n = y_2 - \left(\frac{a_{21}}{a_{11}}\right)y_1$$

$$- -$$

$$0 + \left(a_{n2} - \frac{a_{12}a_{n1}}{a_{11}}\right)x_2 + \cdots + \left(a_{nn} - \frac{a_{1n}a_{n1}}{a_{11}}\right)x_n = y_n - \left(\frac{a_{n1}}{a_{11}}\right)y_1 \qquad (1\text{–}6)$$

Once this has been done we shall have no need for the original values of the a_{ij}'s or the y's. Let us rename the variables in the following way:

$$a_{ij} \quad \text{remains} \quad a_{ij} \quad \text{if} \quad i = 1.$$

$$a_{ij} \quad \text{becomes} \quad 0 \quad \text{if} \quad j = 1, \quad i \neq 1.$$

$$a_{ij} \quad \text{replaces} \quad a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}} \quad \text{if} \quad i \neq 1, \quad j \neq 1.$$

$$y_1 \quad \text{remains} \quad y_1.$$

$$y_j \quad \text{replaces} \quad y_j - \left(\frac{a_{i1}}{a_{11}}\right)y_1 \qquad j \neq 1.$$

Then Eqs. 1–6 become

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1$$

$$0 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2$$

$$- -$$

$$0 + a_{n2}x_2 + \cdots + a_{nn}x_n = y_n \qquad (1\text{–}7)$$

where new parameters are used. We rename the variables using old variable names to reduce the number of variables stored during the computer solution.

Now we proceed in the same way but operate on the second to nth equations. Before doing this, let us introduce some time-saving notation. Instead of writing the x_j's each time, let us use matrix notation. In addition, we shall include the y's in the matrix by adding an $(n + 1)$th column to the matrix. The array shall be called the *augmented matrix* and will be indicated by $\hat{A}'$. Then

$$a_{j,n+1} = y_j, \qquad j = 1, 2, \ldots, n \qquad (1\text{–}8)$$