RANDOM DATA:
ANALYSIS AND MEASUREMENT PROCEDURES

JULIUS S. BENDAT

Mathematical and Environmental Consultant
Formerly President, Measurement Analysis Corporation

ALLAN G. PIERSOL

Staff Consultant, Digitek Corporation
and Lecturer in Engineering, University of Southern California

WILEY-INTERSCIENCE

a division of John Wiley & Sons, Inc.

New York • Chichester • Brisbane • Toronto • Singapore

Library of Congress Catalog Card Number: 71-160211

ISBN 0-471-06470-X

Printed in the United States of America.

20 19 18 17 16 15 14 13

PREFACE

This book is an extensive revision and replacement for the authors' early book, *Measurement and Analysis of Random Data*, 1966. Approximately 50 percent of the original material has been rewritten or deleted and replaced by new material. These changes reflect the technical advances that have taken place in the last five years as well as an increased awareness of pertinent matters gained through the further experience of the authors. Specifically, a broader discussion appears on statistical errors in random data analysis. An entirely new chapter has been introduced to integrate the general requirements for data acquisition, recording, preparation, qualification, and processing. The discussions of digital data analysis procedures have been greatly expanded to cover the more recent analysis techniques made feasible by the availability of fast Fourier transform algorithms. Discussions of transient and multidimensional random processes are now included. Finally, a number of illustrative examples involving actual physical data have been added to support theoretical developments. The illustrations are largely restricted to aerospace and automotive applications since these are the fields of most recent concern to the authors. The general techniques, however, are applicable to data common in many other fields including meteorology, oceanography, seismology, communications, nuclear processes, and biomedical research.

The emphasis in this new book is on the practical aspects of random data analysis and measurement procedures, with special attention to the interrelationships of the various technical disciplines involved. As before, the book is written with the primary intent of providing a convenient reference for practicing engineers and scientists. The secondary intent of providing a specialized textbook for students has been augmented by the addition of problem sets at the end of each chapter. The reader is assumed to have a basic knowledge of probability theory, statistics, and transform methods of applied mathematics.

Summaries of chapter contents appear at the beginning of each chapter. In brief, Chapters 1 through 4 present a review of basic theoretical background material needed for the developments in later chapters. Basic descriptive properties of random data are outlined in Chapter 1 while physical system response properties are reviewed in Chapter 2. Pertinent mathematical and statistical theory is summarized in Chapters 3 and 4. This review material is followed in Chapters 5 and 6 by extensive developments and formulations of input-output relationships and statistical errors in measured data. Chapter 7 outlines the overall procedures for random data acquisition and processing. Detailed procedures for analog and digital data analysis are presented in Chapters 8 and 9. The final Chapter 10 discusses some advanced ideas and procedures relevant to nonstationary, transient, and multidimensional data.

We wish to acknowledge the many contributions to this book by former associates in Measurement Analysis Corporation and Digitek Corporation. We also thank those government agencies, industrial companies, and individuals who supported our work. A special appreciation is given to Engineering Extension, University of California, Los Angeles, and to other organizations, who sponsored our presentation of short courses on this subject matter. Our final thanks extends to Teresia Piersol and Lucinda Bendat for their help in preparing the manuscript.

Los Angeles, California
July 1971

JULIUS S. BENDAT
ALLAN G. PIERSOL

GLOSSARY OF SYMBOLS

a, b	Sample Regression Coefficients, Arbitrary Constants
$b[\ \ ]$	Bias Error of [ ]
B	Cyclical Frequency Bandwidth
c	Mechanical Damping Coefficient, Arbitrary Constant
C	Electrical Capacitance
C_{xy}	Covariance
$C_x(\tau)$	Autocovariance Function
$C_{xy}(\tau)$	Cross-Covariance Function
$C(t_1, t_2)$	Nonstationary Covariance Function
$C_{xy}(f)$	Coincident Spectral Density Function (One-Sided)
D	Time Displacement Range
$e(t)$	Potential Difference
$E[\ \ ]$	Expected Value of [ ]
f	Cyclical Frequency
F	Cyclical Frequency Range, Statistical F Variable
$F(t)$	Mechanical Forcing Function
$G_x(f)$	Power Spectral Density Function Defined for Non-Negative Frequencies Only (One-Sided)
$G_{xy}(f)$	Cross-Spectral Density Function Defined for Non-Negative Frequencies Only (One-Sided)
h	Sampling Interval
$h(\tau)$	Weighting Function (Unit Impulse Response Function)
$H(f)$	Frequency Response Function
$\|H(f)\|$	Gain Factor
$\mathrm{Im}[\ \ ]$	Imaginary Part of [ ]
j	$\sqrt{-1}$, Index
k	Mechanical Spring Constant, Index
K	RC Filter Time Constant, Number of Class Intervals
ℓ	Number of Frequency Components
L	Electrical Inductance, Length
m	Mechanical Mass, Maximum Number of Lag Values
m_f	Modulation Index
n	Degrees-of-Freedom
N	Sample Size
$p(x)$	Probability Density Function

$p(x, y)$	Joint Probability Density Function
$P(x)$	Probability Distribution Function
$P(x, y)$	Joint Probability Distribution Function
Prob[ ]	Probability that [ ]
q	Number of Inputs, Number of Sample Records
$q(t)$	Electrical Charge
$Q_{xy}(f)$	Quadrature Spectral Density Function (One-Sided)
r	Number of Runs, Arbitrary Factor
r_{xy}	Sample Correlation Coefficient
R	Electrical Resistance
R_s	Analyzer Scan Rate
$R_x(\tau)$	Autocorrelation Function
$R_{xy}(\tau)$	Cross-Correlation Function
$R(t_1, t_2)$	Nonstationary Correlation Function
Re [ ]	Real Part of [ ]
s	Sample Standard Deviation
s^2	Sample Variance
${s_{xy}}^2$	Sample Covariance
$S_x(f)$	Power Spectral Density Function Defined for Both Positive and Negative Frequencies (Two-Sided)
$S_{xy}(f)$	Cross-Spectral Density Function Defined for Both Positive and Negative Frequencies (Two-Sided)
$S(f_1, f_2)$	Generalized (Nonstationary) Spectral Density Function
S/N	Signal to Noise Ratio
t	Time Variable, Student t Variable
T	Observation Time, Averaging Time
T_s	Analysis Time
u_n	Raw Data Values
$u(x, t)$	Space and Time Dependent Variable
V	Voltage Range
Var [ ]	Variance of [ ]
W	Amplitude Window Width
$x(t)$, $y(t)$	Time Dependent Variables
$\bar{x}$	Sample Mean Value of x
$\overline{\lvert x \rvert}$	Mean Absolute Value (Average Rectified Value) of x
X	Amplitude of Sinusoidal $x(t)$
$X(f)$	Fourier Transform of $x(t)$
z	Standardized Normal Variable
$\lvert[\ ]\rvert$	Absolute Value of [ ]
$\hat{[\ ]}$	Estimate of [ ]
α	A Small Probability, Level of Significance
β	Probability of a Type II error

$\gamma^2(f)$	Coherence Function
$\delta(\ )$	Delta Function
Δ	Small Increment
ϵ	Normalized Error
ζ	Mechanical Damping Ratio
θ	Phase Angle
$\theta_{xy}(f)$	Argument of $G_{xy}(f)$
μ	Mean Value
ρ	Correlation Coefficient
$\rho(\tau)$	Correlation Function Coefficient
σ	Standard Deviation
σ^2	Variance
τ	Time Displacement
$\phi(f)$	Phase Factor
Φ	Arbitrary Statistical Parameter
χ^2	Statistical Chi-Square Variable
Ψ	Root Mean Square Value
Ψ^2	Mean Square Value
λ	Wavelength

CONTENTS

CHAPTER 1

BASIC DESCRIPTIONS OF PHYSICAL DATA

Any observed data representing a physical phenomenon can be broadly classified as being either deterministic or nondeterministic. Deterministic data are those that can be described by an explicit mathematical relationship. For example, consider a rigid body which is suspended from a fixed foundation by a linear spring, as shown in Figure 1.1. Let m be the mass of the body (assumed to be inelastic) and k be the spring constant of the spring (assumed to be massless). Suppose the body is displaced from its position of equilibrium by a distance X, and released at time $t = 0$. From either basic laws of mechanics or repeated observations, it can be established that the following relationship will apply.

$$x(t) = X \cos \sqrt{\frac{k}{m}}\, t \qquad t \geq 0 \tag{1.1}$$

Equation (1.1) defines the exact location of the body at any instant of time in the future. Hence the physical data representing the motion of the mass are deterministic.

There are many physical phenomena in practice which produce data that can be represented with reasonable accuracy by explicit mathematical relationships. For example, the motion of a satellite in orbit about the earth, the potential across a condenser as it discharges through a resistor, the vibration response of an unbalanced rotating machine, or the temperature of water as heat is applied, are all basically deterministic. However, there are many other physical phenomena which produce data that are not deterministic. For example, the height of waves in a confused sea, the acoustic pressures generated by air rushing through a pipe, or the electrical

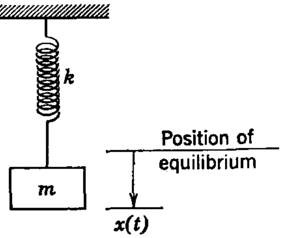

Figure 1.1 Simple spring mass system.

output of a noise generator represent data which cannot be described by explicit mathematical relationships. There is no way to predict an exact value at a future instant of time. These data are random in character and must be described in terms of probability statements and statistical averages rather than by explicit equations.

The classification of various physical data as being either deterministic or random might be debated in many cases. For example, it might be argued that no physical data in practice can be truly deterministic since there is always a possibility that some unforeseen event in the future might influence the phenomenon producing the data in a manner that was not originally considered. On the other hand, it might be argued that no physical data are truly random since exact mathematical descriptions might be possible if a sufficient knowledge of the basic mechanisms of the phenomenon producing the data were known. In practical terms, the decision as to whether or not physical data are deterministic or random is usually based upon the ability to reproduce the data by controlled experiments. If an experiment producing specific data of interest can be repeated many times with identical results (within the limits of experimental error), then the data can generally be considered deterministic. If an experiment cannot be designed which will produce identical results when the experiment is repeated, then the data must usually be considered random in nature.

Various special classifications of deterministic and random data will now be discussed. Note that the classifications are selected from an analysis viewpoint and do not necessarily represent the most suitable classifications from other possible viewpoints. Further note that physical data are usually thought of as being functions of time and will be discussed in such terms for convenience. However, any other variable can replace time as required.

1.1 CLASSIFICATIONS OF DETERMINISTIC DATA

Data representing deterministic phenomena can be categorized as being either periodic or nonperiodic. Periodic data can be further categorized as

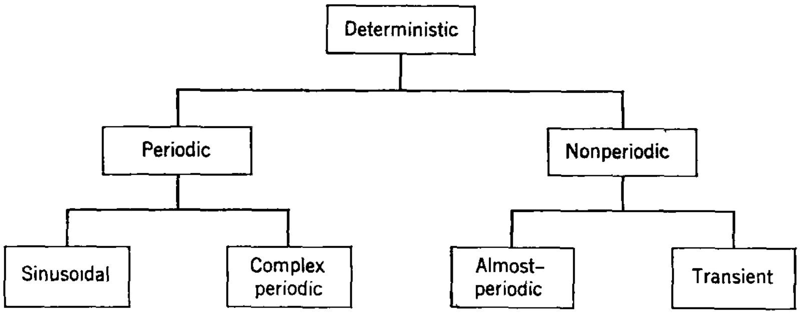

Figure 1.2 Classifications of deterministic data.

being either sinusoidal or complex periodic. Nonperiodic data can be further categorized as being either "almost-periodic" or transient. These various classifications of deterministic data are schematically illustrated in Figure 1.2. Of course, any combination of these forms may also occur. For purposes of review, each of these types of deterministic data along with physical examples will be briefly discussed.

1.1.1 *Sinusoidal Periodic Data*

Sinusoidal data are those types of periodic data which can be defined mathematically by a time-varying function of the form

$$x(t) = X \sin (2\pi f_0 t + \theta) \tag{1.2}$$

where X = amplitude

f_0 = cyclical frequency in cycles per unit time

θ = initial phase angle with respect to the time origin in radians

$x(t)$ = instantaneous value at time t

The sinusoidal time history described by Equation (1.2) is usually referred to as a sine wave. When analyzing sinusoidal data in practice, the phase angle θ is often ignored. For this case

$$x(t) = X \sin 2\pi f_0 t \tag{1.3}$$

Equation (1.3) can be pictured by a time history plot or by an amplitude-frequency plot (frequency spectrum), as illustrated in Figure 1.3.

The time interval required for one full fluctuation or cycle of sinusoidal data is called the period T_p. The number of cycles per unit time is called the frequency f_0. The frequency and period are related by

$$T_p = \frac{1}{f_0} \tag{1.4}$$

Note that the frequency spectrum in Figure 1.3 is composed of an amplitude component at a specific frequency, as opposed to a continuous plot of amplitude versus frequency. Such spectra are called *discrete spectra* or *line spectra*.

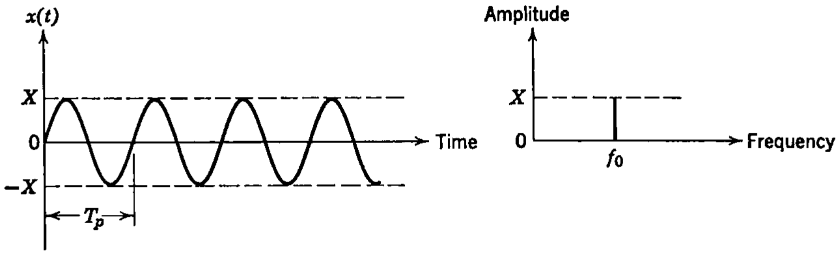

Figure 1.3 Time history and spectrum of sinusoidal data.

There are many examples of physical phenomena which produce approximately sinusoidal data in practice. The voltage output of an electrical alternator is one example; the vibratory motion of an unbalanced rotating weight is another. Sinusoidal data represent one of the simplest forms of time-varying data from the analysis viewpoint.

1.1.2 *Complex Periodic Data*

Complex periodic data are those types of periodic data which can be defined mathematically by a time-varying function whose waveform exactly repeats itself at regular intervals such that

$$x(t) = x(t \pm nT_p) \qquad n = 1, 2, 3, \ldots \tag{1.5}$$

As for sinusoidal data, the time interval required for one full fluctuation is called the *period* T_p. The number of cycles per unit time is called the *fundamental frequency* f_1. A special case for complex periodic data is clearly sinusoidal data where $f_1 = f_0$.

With few exceptions in practice, complex periodic data may be expanded into a Fourier series according to the following formula.

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos 2\pi n f_1 t + b_n \sin 2\pi n f_1 t) \tag{1.6}$$