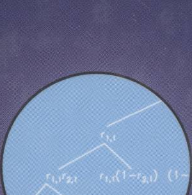
WILEY

MULTISCALE ANALYSIS OF COMPLEX TIME SERIES

Integration of Chaos and Random Fractal Theory, and Beyond



WWW. LINK AVAILABLE

JIANBO GAO • YINHE CAO • WEN-WEN TUNG • JING HU

MULTISCALE ANALYSIS OF COMPLEX TIME SERIES

Integration of Chaos and Random Fractal Theory, and Beyond

Jianbo Gao
Department of Electrical and Computer Engineering
University of Florida

Yinhe Cao
BioSieve

Wen-wen Tung
Department of Earth and Atmospheric Sciences
Purdue University

Jing Hu
Department of Electrical and Computer Engineering
University of Florida

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Multiscale analysis of complex time series : integration of chaos and random
fractal theory, and beyond / Jianbo Gao . . . [et al.].
p. cm.
Includes index.
ISBN 978-0-471-65470-4 (cloth)
1. Time series analysis. 2. Chaotic behavior in systems. 3. Fractals. I.
Gao, Jianbo, 1966–
QA280.M85 2007
519.5'5—dc22 2007019072

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

MULTISCALE ANALYSIS OF COMPLEX TIME SERIES

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

THE WILEY BICENTENNIAL–KNOWLEDGE FOR GENERATIONS

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

WILLIAM J. PESCE
PRESIDENT AND CHIEF EXECUTIVE OFFICER

PETER BOOTH WILEY
CHAIRMAN OF THE BOARD

To our teachers,
Xianyi Zhou of Zhejiang
University,
Yilong Bai and Zhemin
Zheng of Chinese Academy
of Sciences,
Michio Yanai of UCLA,
and to our families

PREFACE

Complex interconnected systems, including the Internet, stock markets, and human heart or brain, are usually comprised of multiple subsystems that exhibit highly nonlinear deterministic as well as stochastic characteristics and are regulated hierarchically. They generate signals that exhibit complex characteristics such as nonlinearity, sensitive dependence on small disturbances, long memory, extreme variations, and nonstationarity. A complex system usually cannot be studied by decomposing the system into its constituent subsystems, but rather by measuring certain signals generated by the system and analyzing the signals to gain insights into the behavior of the system. In this endeavor, data analysis is a crucial step. Chaos theory and random fractal theory are two of the most important theories developed for data analysis. Unfortunately, no single book has been available to present all the basic concepts necessary for researchers to fully understand the ever-expanding literature and apply novel methods to effectively solve their signal processing problems. This book attempts to meet this pressing need by presenting chaos theory and random fractal theory in a unified way.

Integrating chaos theory and random fractal theory and going beyond them has proven to be much harder than we had thought, because the foundations for chaos theory and random fractal theory are entirely different. Chaos theory is mainly concerned about apparently irregular behaviors in a complex system that are generated by nonlinear deterministic interactions of only a few numbers of degrees of freedom, where noise or intrinsic randomness does not play an important role, while random fractal theory assumes that the dynamics of the system are inherently random. After postponing delivery of the book for more than two and half years, we are finally satisfied. The book now contains many new results in Chapters 8–15 that have not been published elsewhere, culminating in the development of a multiscale complexity measure that is computable from short, noisy time series. As shown in Chapter 15, the measure can readily classify major types of complex motions, effectively deal with nonstationarity, and simultaneously characterize the behaviors of complex signals on a wide range of scales, including complex irregular behaviors on small scales and orderly behaviors, such as oscillatory motions, on large scales.

This book has adopted a data-driven approach. To help readers better understand and appreciate the power of the materials in the book, nearly every significant concept or approach presented is illustrated by applying it to effectively solve real problems, sometimes with unprecedented accuracy. Furthermore, source codes, written in various languages, including Java, Fortran, C, and Matlab, for many methods are provided in a dedicated book website, together with some simulated and experimental data (see Sec. A.4 in Appendix A).

This book contains enough material for a one-year graduate-level course. It is useful for students with various majors, including electrical engineering, computer science, civil and environmental engineering, mechanical engineering, chemical engineering, medicine, chemistry, physics, geophysics, mathematics, finance, and population ecology. It is also useful for researchers working in relevant fields and practitioners who have to solve their own signal processing problems.

We thank Drs. Vince Billock, Gijs Bosman, Yenn-Ru Chen, Yuguang Fang, Jose Fortes, John Harris, Hsiao-ming Hsu, Sheng-Kwang Hwang, Mark Law, Jian Li, Johnny Lin, Jiamin Liu, Mitch Moncrieff, Jose Principe, Vladimir Protopopescu, Nageswara Rao, Ronn Ritke, Vwani Roychowdhury, Izhak Rubin, Chris Sackellares, Zhen-Su She, Yuch-Ning Shieh, Peter Stoica, Martin Uman, Kung Yao, and Keith White for many useful discussions. Drs. Jon Harbor, Andy Majda, and Robert Nowack have read part of Chapter 15, while Dr. Alexandre Chorin has read a number of chapters. We are grateful for their many useful suggestions and encouragement. One of the authors, Jianbo Gao, taught a one-year course entitled "Signal Processing with Chaos and Fractals" at the University of Florida, in the fall of 2002 and the spring of 2003. Students' enthusiasm has been instrumental in driving us to finish the book. He would particularly thank his former and current students Jing Ai, Ung Sik Kim, Jaemin Lee, Yan Qi, Dongming Xu, and Yi Zheng for their contributions to the many topics presented here. We would like to thank the editors at Wiley, Helen Greenberg, Whitney Lesch, Val Moliere, Christine Punzo, and George Telecki, for their patience and encouragement. Finally, we thank IPAM at UCLA and MBI at the Ohio State University for generously supporting us to attend a number of interesting workshops organized by the two institutions.

CONTENTS

Preface xiii

1 Introduction **1**

1.1 Examples of multiscale phenomena 4
1.2 Examples of challenging problems to be pursued 9
1.3 Outline of the book 12
1.4 Bibliographic notes 14

2 Overview of fractal and chaos theories **15**

2.1 Prelude to fractal geometry 15
2.2 Prelude to chaos theory 18
2.3 Bibliographic notes 23
2.4 Warmup exercises 23

3 Basics of probability theory and stochastic processes **25**

3.1 Basic elements of probability theory 25
3.1.1 Probability system 25
3.1.2 Random variables 27
3.1.3 Expectation 30

3.1.4 Characteristic function, moment generating function, Laplace transform, and probability generating function 32
3.2 Commonly used distributions 34
3.3 Stochastic processes 41
3.3.1 Basic definitions 41
3.3.2 Markov processes 43
3.4 Special topic: How to find relevant information for a new field quickly 49
3.5 Bibliographic notes 51
3.6 Exercises 51

4 Fourier analysis and wavelet multiresolution analysis 53

4.1 Fourier analysis 54
4.1.1 Continuous-time (CT) signals 54
4.1.2 Discrete-time (DT) signals 55
4.1.3 Sampling theorem 57
4.1.4 Discrete Fourier transform 58
4.1.5 Fourier analysis of real data 58
4.2 Wavelet multiresolution analysis 62
4.3 Bibliographic notes 67
4.4 Exercises 67

5 Basics of fractal geometry 69

5.1 The notion of dimension 69
5.2 Geometrical fractals 71
5.2.1 Cantor sets 71
5.2.2 Von Koch curves 74
5.3 Power law and perception of self-similarity 75
5.4 Bibliographic notes 76
5.5 Exercises 76

6 Self-similar stochastic processes 79

6.1 General definition 79
6.2 Brownian motion (Bm) 81
6.3 Fractional Brownian motion (fBm) 84
6.4 Dimensions of Bm and fBm processes 87
6.5 Wavelet representation of fBm processes 89
6.6 Synthesis of fBm processes 90
6.7 Applications 93
6.7.1 Network traffic modeling 93
6.7.2 Modeling of rough surfaces 97
6.8 Bibliographic notes 97

6.9 Exercises 98

7 Stable laws and Levy motions 99

7.1 Stable distributions 100
7.2 Summation of strictly stable random variables 103
7.3 Tail probabilities and extreme events 104
7.4 Generalized central limit theorem 107
7.5 Levy motions 108
7.6 Simulation of stable random variables 109
7.7 Bibliographic notes 111
7.8 Exercises 112

8 Long memory processes and structure-function–based multifractal analysis 115

8.1 Long memory: basic definitions 115
8.2 Estimation of the Hurst parameter 118
8.3 Random walk representation and structure-function–based multifractal analysis 119
8.3.1 Random walk representation 119
8.3.2 Structure-function–based multifractal analysis 120
8.3.3 Understanding the Hurst parameter through multifractal analysis 121
8.4 Other random walk–based scaling parameter estimation 124
8.5 Other formulations of multifractal analysis 124
8.6 The notion of finite scaling and consistency of H estimators 126
8.7 Correlation structure of ON/OFF intermittency and Levy motions 130
8.7.1 Correlation structure of ON/OFF intermittency 130
8.7.2 Correlation structure of Levy motions 131
8.8 Dimension reduction of fractal processes using principal component analysis 132
8.9 Broad applications 137
8.9.1 Detection of low observable targets within sea clutter 137
8.9.2 Deciphering the causal relation between neural inputs and movements by analyzing neuronal firings 139
8.9.3 Protein coding region identification 147
8.10 Bibliographic notes 149
8.11 Exercises 151

9 Multiplicative multifractals 153

9.1 Definition 153
9.2 Construction of multiplicative multifractals 154
9.3 Properties of multiplicative multifractals 157

9.4 Intermittency in fully developed turbulence 163
9.4.1 Extended self-similarity 165
9.4.2 The log-normal model 167
9.4.3 The log-stable model 168
9.4.4 The β-model 168
9.4.5 The random β-model 168
9.4.6 The p model 169
9.4.7 The SL model and log-Poisson statistics of turbulence 169
9.5 Applications 171
9.5.1 Target detection within sea clutter 173
9.5.2 Modeling and discrimination of human neuronal activity 173
9.5.3 Analysis and modeling of network traffic 176
9.6 Bibliographic notes 178
9.7 Exercises 179

10 Stage-dependent multiplicative processes **181**

10.1 Description of the model 181
10.2 Cascade representation of $1/f^{\beta}$ processes 184
10.3 Application: Modeling heterogeneous Internet traffic 189
10.3.1 General considerations 189
10.3.2 An example 191
10.4 Bibliographic notes 193
10.5 Exercises 193

11 Models of power-law-type behavior **195**

11.1 Models for heavy-tailed distribution 195
11.1.1 Power law through queuing 195
11.1.2 Power law through approximation by log-normal distribution 196
11.1.3 Power law through transformation of exponential distribution 197
11.1.4 Power law through maximization of Tsallis nonextensive entropy 200
11.1.5 Power law through optimization 202
11.2 Models for $1/f^{\beta}$ processes 203
11.2.1 $1/f^{\beta}$ processes from superposition of relaxation processes 203
11.2.2 $1/f^{\beta}$ processes modeled by ON/OFF trains 205
11.2.3 $1/f^{\beta}$ processes modeled by self-organized criticality 206
11.3 Applications 207
11.3.1 Mechanism for long-range-dependent network traffic 207
11.3.2 Distributional analysis of sea clutter 209
11.4 Bibliographic notes 210
11.5 Exercises 211

12 Bifurcation theory **213**

12.1 Bifurcations from a steady solution in continuous time systems 213
12.1.1 General considerations 214
12.1.2 Saddle-node bifurcation 215
12.1.3 Transcritical bifurcation 215
12.1.4 Pitchfork bifurcation 215
12.2 Bifurcations from a steady solution in discrete maps 217
12.3 Bifurcations in high-dimensional space 218
12.4 Bifurcations and fundamental error bounds for fault-tolerant computations 218
12.4.1 Error threshold values for arbitrary K-input NAND gates 219
12.4.2 Noisy majority gate 222
12.4.3 Analysis of von Neumann's multiplexing system 226
12.5 Bibliographic notes 233
12.6 Exercises 233

13 Chaotic time series analysis **235**

13.1 Phase space reconstruction by time delay embedding 236
13.1.1 General considerations 236
13.1.2 Defending against network intrusions and worms 237
13.1.3 Optimal embedding 240
13.2 Characterization of chaotic attractors 243
13.2.1 Dimension 244
13.2.2 Lyapunov exponents 246
13.2.3 Entropy 251
13.3 Test for low-dimensional chaos 254
13.4 The importance of the concept of scale 258
13.5 Bibliographic notes 258
13.6 Exercises 259

14 Power-law sensitivity to initial conditions (PSIC) **261**

14.1 Extending exponential sensitivity to initial conditions to PSIC 262
14.2 Characterizing random fractals by PSIC 263
14.2.1 Characterizing $1/f^{\beta}$ processes by PSIC 264
14.2.2 Characterizing Levy processes by PSIC 265
14.3 Characterizing the edge of chaos by PSIC 266
14.4 Bibliographic notes 268

15 Multiscale analysis by the scale-dependent Lyapunov exponent (SDLE) **271**

15.1 Basic theory 271
15.2 Classification of complex motions 274

15.2.1 Chaos, noisy chaos, and noise-induced chaos 274
15.2.2 $1/f^{\beta}$ processes 276
15.2.3 Levy flights 277
15.2.4 SDLE for processes defined by PSIC 279
15.2.5 Stochastic oscillations 279
15.2.6 Complex motions with multiple scaling behaviors 280
15.3 Distinguishing chaos from noise 283
15.3.1 General considerations 283
15.3.2 A practical solution 284
15.4 Characterizing hidden frequencies 286
15.5 Coping with nonstationarity 290
15.6 Relation between SDLE and other complexity measures 291
15.7 Broad applications 297
15.7.1 EEG analysis 297
15.7.2 HRV analysis 298
15.7.3 Economic time series analysis 300
15.7.4 Sea clutter modeling 303
15.8 Bibliographic notes 304

Appendix A: Description of data **307**

A.1 Network traffic data 307
A.2 Sea clutter data 308
A.3 Neuronal firing data 309
A.4 Other data and program listings 309

Appendix B: Principal Component Analysis (PCA), Singular Value Decomposition (SVD), and Karhunen-Loève (KL) expansion **311**

Appendix C: Complexity measures **313**

C.1 FSLE 314
C.2 LZ complexity 315
C.3 PE 317

References 319

Index 347

CHAPTER 1

INTRODUCTION

Complex systems are usually comprised of multiple subsystems that exhibit both highly nonlinear deterministic and stochastic characteristics and are regulated hierarchically. These systems generate signals that exhibit complex characteristics such as sensitive dependence on small disturbances, long memory, extreme variations, and nonstationarity. A stock market, for example, is strongly influenced by multilayered decisions made by market makers, as well as by interactions of heterogeneous traders, including intraday traders, short-period traders, and long-period traders, and thus gives rise to highly irregular stock prices. The Internet, as another example, has been designed in a fundamentally decentralized fashion and consists of a complex web of servers and routers that cannot be effectively controlled or analyzed by traditional tools of queuing theory or control theory and give rise to highly bursty and multiscale traffic with extremely high variance, as well as complex dynamics with both deterministic and stochastic components. Similarly, biological systems, being heterogeneous, massively distributed, and highly complicated, often generate nonstationary and multiscale signals. With the rapid accumulation of complex data in health sciences, systems biology, nano-sciences, information systems, and physical sciences, it has become increasingly important to be able to analyze multiscale and nonstationary data.

Multiscale signals behave differently, depending upon the scale at which the data are examined. How can the behaviors of such signals on a wide range of scales be simultaneously characterized? One strategy we envision is to use existing theories synergistically instead of individually. To make this possible, appropriate scale ranges where each theory is most pertinent need to be identified. This is a difficult task, however, since different theories may have entirely different foundations. For example, chaos theory is mainly concerned about apparently irregular behaviors in a complex system that are generated by nonlinear deterministic interactions with only a few degrees of freedom, where noise or intrinsic randomness does not play an important role. Random fractal theory, on the other hand, assumes that the dynamics of the system are inherently random. Therefore, to make this strategy work, different theories need to be integrated and even generalized.

The second vital strategy we envision is to develop measures that explicitly incorporate the concept of scale so that different behaviors of the data on varying scales can be simultaneously characterized by the same scale-dependent measure. In the most ideal scenario, a scale-dependent measure can readily classify different types of motions based on analysis of short, noisy data. In this case, one can readily see that the measure will be able not only to identify appropriate scale ranges where different theories, including information theory, chaos theory, and random fractal theory, are applicable, but also to automatically characterize the behaviors of the data on those scale ranges.

The vision presented above dictates the style and the scope of this book, as depicted in Fig. 1.1. Specifically, we aim to build an effective arsenal by synergistically integrating approaches based on chaos and random fractal theory, and going beyond this, to complement conventional approaches such as spectral analysis and machine learning techniques. To make such an integration possible, four important efforts are made:

1. Wavelet representation of fractal models as well as wavelet estimation of fractal scaling parameters will be carefully developed. Furthermore, a new fractal model will be developed. The model provides a new means of characterizing long-range correlations in time series and a convenient way of modeling non-Gaussian statistics. More importantly, it ties together different approaches in the vast field of random fractal theory (represented by the four small boxes under the "Random Fractal" box in Fig. 1.1).

2. Fractal scaling break and truncation of power-law behavior are related to specific features of real data so that scale-free fractal behavior as well as structures defined by specific scales can be simultaneously characterized.

3. A new theoretical framework for signal processing — power-law sensitivity to initial conditions (PSIC) — will be developed, to provide chaos and random fractal theory a common foundation so that they can be better integrated.

4. The scale-dependent Lyapunov exponent (SDLE), which is a variant of the finite-size Lyapunov exponent (FSLE), is an excellent multiscale measure. We shall develop a highly efficient algorithm for calculating it and show that it can readily classify different types of motions, aptly characterize complex behaviors of real-world multiscale signals on a wide range of scales, and, therefore, naturally solve the classic problem of distinguishing low-dimensional chaos from noise. Furthermore, we shall show that the SDLE can effectively deal with nonstationarity and that existing complexity measures can be related to the value of the SDLE on specific scales.

To help readers better understand and appreciate the power of the materials in this book, nearly every significant concept or approach presented will be illustrated by applying it to effectively solve real problems, sometimes with unprecedented accuracy. Furthermore, source codes, written in various languages, including Fortran, C, and Matlab, for many methods are provided together with some simulated and experimental data.

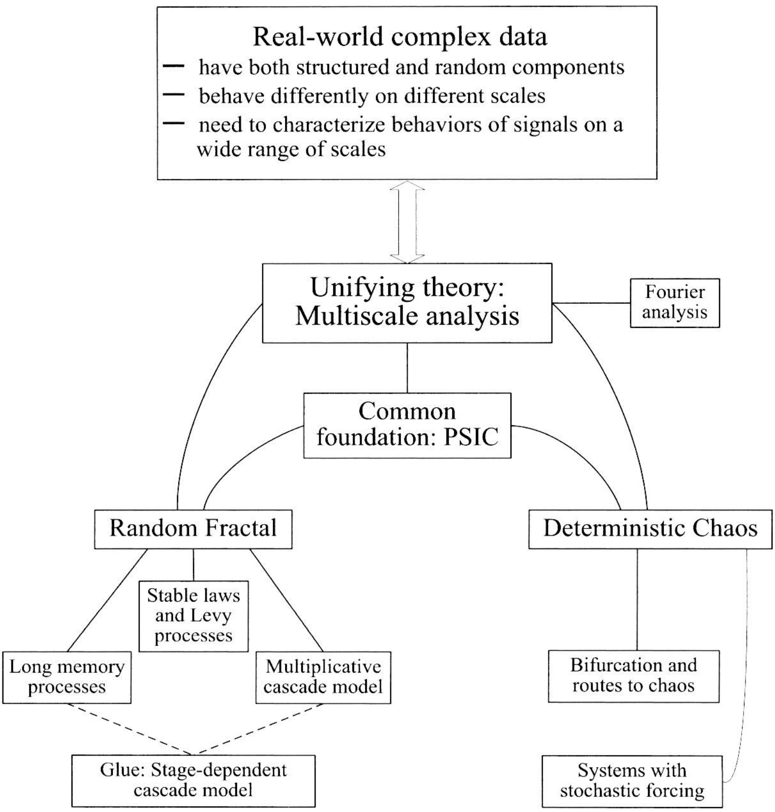


Figure 1.1. Structure of the book.