

AN INTRODUCTION TO SYMBOLIC DYNAMICS AND CODING

Second Edition

Symbolic dynamics is a mature yet rapidly developing area of dynamical systems. It has established strong connections with many areas, including linear algebra, graph theory, probability, group theory, and the theory of computation, as well as data storage, statistical mechanics, and C^* -algebras.

This Second Edition maintains the introductory character of the original 1995 edition as a general textbook on symbolic dynamics and its applications to coding. It is written at an elementary level and aimed at students, well-established researchers, and experts in mathematics, electrical engineering, and computer science. Topics are carefully developed and motivated with many illustrative examples. There are more than 500 exercises to test the reader's understanding. In addition to a chapter in the First Edition on advanced topics and a comprehensive bibliography, the Second Edition includes a detailed Addendum, with companion bibliography, describing major developments and new research directions since publication of the First Edition.

Douglas Lind is Professor Emeritus of Mathematics at the University of Washington. He was department chair, is an Inaugural Fellow of the American Mathematical Society, and served in many governance roles for the Mathematical Sciences Research Institute, including chairing the committee that designed the 2006 addition.

Brian Marcus is Professor of Mathematics at the University of British Columbia. He shared the 1993 Leonard Abraham Prize Paper award of the IEEE Communications Society. He is currently the UBC Site Director of the Pacific Institute for the Mathematical Sciences (PIMS) and is a Fellow of the AMS and IEEE.



CAMBRIDGE MATHEMATICAL LIBRARY

Cambridge University Press has a long and honourable history of publishing in mathematics and counts many classics of the mathematical literature within its list. Some of these titles have been out of print for many years now and yet the methods which they espouse are still of considerable relevance today.

The *Cambridge Mathematical Library* provides an inexpensive edition of these titles in a durable paperback format and at a price that will make the books attractive to individuals wishing to add them to their own personal libraries. Certain volumes in the series have a foreword, written by a leading expert in the subject, which places the title in its historical and mathematical context.

A complete list of books in the series can be found at www.cambridge.org/mathematics. Recent titles include the following:

Reversibility and Stochastic Networks F. P. KELLY

The Geometry of Moduli Spaces of Sheaves (Second Edition)
DANIEL HUYBRECHTS & MANFRED LEHN

Smooth Compactifications of Locally Symmetric Varieties (Second Edition)
AVNER ASH, DAVID MUMFORD, MICHAEL RAPOPORT &
YUNG-SHENG TAI

Markov Chains and Stochastic Stability (Second Edition) SEAN MEYN & RICHARD L. TWEEDIE

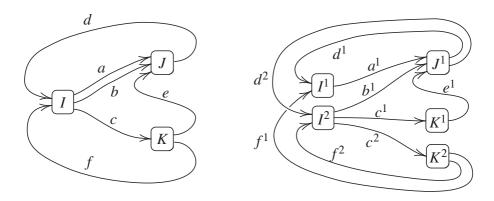
Combinatorics: The Rota Way
JOSEPH P. S. KUNG, GIAN-CARLO ROTA & CATHERINE H. YAN

A Course of Pure Mathematics (Centenary Edition) G. H. HARDY

Weather Prediction by Numerical Process (Second Edition)
LEWIS FRY RICHARDSON

Spline Functions: Basic Theory (Third Edition)
LARRY SCHUMAKER





An example of state splitting, a fundamental operation in symbolic dynamics. By partitioning outgoing edges, the states of the graph on the left are split to form the graph on the right.



AN INTRODUCTION TO SYMBOLIC DYNAMICS AND CODING

Second Edition

DOUGLAS LIND
University of Washington

BRIAN MARCUS

University of British Columbia







Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

 ${\bf www.cambridge.org}$ Information on this title: www.cambridge.org/9781108820288

DOI: 10.1017/9781108899727

© Cambridge University Press & Assessment 1995, 2021

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 1995 Second edition 2021

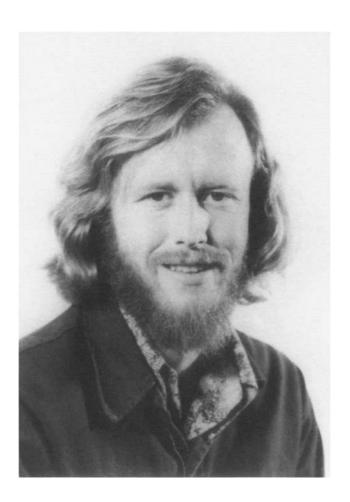
A catalogue record for this publication is available from the British Library

 ${\tt ISBN-978\text{-}1\text{-}108\text{-}82028\text{-}8-Paperback}$

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.



Dedicated to the memory of Rufus Bowen, our friend and teacher 1947–1978





CONTENTS

PREFACE TO THE FIRST EDITION	xiii
PREFACE TO THE SECOND EDITION	xix
CHAPTER 1. SHIFT SPACES	. 1
§1.1. Full Shifts	. 1
§1.2. Shift Spaces	
§1.3. Languages	
§1.4. Higher Block Shifts and Higher Power Shifts	
§1.5. Sliding Block Codes	
§1.6. Convolutional Encoders	
CHAPTER 2. SHIFTS OF FINITE TYPE	28
§2.1. Finite Type Constraints	28
§2.2. Graphs and Their Shifts	33
§2.3. Graph Representations of Shifts of Finite Type	41
§2.4. State Splitting	49
§2.5. Data Storage and Shifts of Finite Type	59
CHAPTER 3. SOFIC SHIFTS	64
§3.1. Presentations of Sofic Shifts	64
§3.2. Characterizations of Sofic Shifts	70
§3.3. Minimal Right-Resolving Presentations	76
§3.4. Constructions and Algorithms	86
CHAPTER 4. ENTROPY	100
§4.1. Definition and Basic Properties	100
§4.2. Perron–Frobenius Theory	
§4.3. Computing Entropy	113
§4.4. Irreducible Components	118
§4.5. Cyclic Structure	126
CHAPTER 5. FINITE-STATE CODES	137
§5.1. Road Colorings and Right-Closing Labelings	138
§5.2. Finite-State Codes	
§5.3. Approximate Eigenvectors	
§5.4. Code Construction	
§5.5. Sliding Block Decoders	



x Contents

CHAPTER 6. SHIFTS AS DYNAMICAL SYSTEMS .	172
§6.1. Metric Spaces	173
§6.2. Dynamical Systems	
§6.3. Invariants	
§6.4. Zeta Functions	
§6.5. Markov Partitions	
CHAPTER 7. CONJUGACY	217
§7.1. The Decomposition Theorem	
§7.2. Strong Shift Equivalence	
§7.3. Shift Equivalence	
§7.4. Invariants for Shift Equivalence	
§7.5. Shift Equivalence and the Dimension Group	
CHAPTER 8. FINITE-TO-ONE CODES AND FINITE	
EQUIVALENCE	265
§8.1. Finite-to-One Codes	265
§8.2. Right-Resolving Codes	
§8.3. Finite Equivalence	
§8.4. Right-Resolving Finite Equivalence	
CHAPTER 9. DEGREES OF CODES AND ALMOST	
CONJUGACY	302
§9.1. The Degree of a Finite-to-One Code	302
§9.2. Almost Invertible Codes	
§9.3. Almost Conjugacy	
§9.4. Typical Points According to Probability	
CHAPTER 10. EMBEDDINGS AND FACTOR CODES	338
§10.1. The Embedding Theorem	338
§10.2. The Masking Lemma	355
§10.3. Lower Entropy Factor Codes	
CHAPTER 11. REALIZATION	
§11.1. Realization of Entropies	370
§11.2. Realization of Zeta Functions	385
$\S 11.3$. Pure Subgroups of Dimension Groups	397
CHAPTER 12. EQUAL ENTROPY FACTORS	402
§12.1. Right-Closing Factors	403
§12.2. Eventual Factors of Equal Entropy	
§12.3. Ideal Classes	
912.3. Ideal Classes	



> Contents хi CHAPTER 13. GUIDE TO ADVANCED TOPICS 430 §13.1. More on Shifts of Finite Type and Sofic Shifts 430 §13.2. Automorphisms of Shifts of Finite Type §13.3. Symbolic Dynamics and Stationary Processes 440 §13.4. Symbolic Dynamics and Ergodic Theory 444 §13.6. Continuous Flows §13.7. Minimal Shifts §13.9. Shifts with a Countable Alphabet §13.10. Higher Dimensional Shifts ADDENDUM §A.1. Classification Problems §A.2. Factor Codes and Embeddings §A.3. Symbolic Models for Smooth Systems §A.5. Automorphism Groups of Shifts 481 §A.8. Symbolic Dynamics over Countable Groups 502 §A.9. Symbolic Representations of Algebraic Actions 508 NOTATION INDEX

INDEX . . .



PREFACE TO FIRST EDITION

Symbolic dynamics is a rapidly growing part of dynamical systems. Although it originated as a method to study general dynamical systems, the techniques and ideas have found significant applications in data storage and transmission as well as linear algebra. This is the first general textbook on symbolic dynamics and its applications to coding, and we hope that it will stimulate both engineers and mathematicians to learn and appreciate the subject.

Dynamical systems originally arose in the study of systems of differential equations used to model physical phenomena. The motions of the planets, or of mechanical systems, or of molecules in a gas can be modeled by such systems. One simplification in this study is to discretize time, so that the state of the system is observed only at discrete ticks of a clock, like a motion picture. This leads to the study of the iterates of a single transformation. One is interested in both quantitative behavior, such as the average time spent in a certain region, and also qualitative behavior, such as whether a state eventually becomes periodic or tends to infinity. Symbolic dynamics arose as an attempt to study such systems by means of discretizing space as well as time. The basic idea is to divide up the set of possible states into a finite number of pieces, and keep track of which piece the state of the system lies in at every tick of the clock. Each piece is associated with a "symbol," and in this way the evolution of the system is described by an infinite sequence of symbols. This leads to a symbolic dynamical system that mirrors and helps us to understand the dynamical behavior of the original system.

In their fundamental paper [MorH1] written over fifty years ago, Morse and Hedlund named the subject of symbolic dynamics and described its philosophy as follows.

The methods used in the study of recurrence and transitivity frequently combine classical differential analysis with a more abstract symbolic analysis. This involves a characterization of the ordinary dynamical trajectory by an unending sequence of symbols termed a symbolic trajectory such that the properties of recurrence and transitivity of the dynamical trajectory are reflected in analogous properties of its symbolic trajectory.

One example of this idea that you are very familiar with is the decimal expansion of real numbers in the unit interval [0,1). Here the transformation is given by multiplying a number $x \in [0,1)$ by 10 and keeping the its fractional part $\{10x\}$. We partition [0,1) into ten equal subintervals [0,1/10), [1/10,2/10), ..., [9/10,1), and we use the "symbol" j for the interval [j/10,(j+1)/10). Let x be a number in [0,1). Using the transformation $x \mapsto \{10x\}$ and this partition, we write down an infinite sequence



xiv

Preface to first edition

of symbols corresponding to any given $x = .x_1x_2...$ as follows. Since $x \in [x_1/10, (x_1+1)/10)$, the first symbol is just the first digit x_1 of its decimal expansion. Since $\{10x\} \in [x_2/10, (x_2+1)/10)$, the next symbol is the first digit of the decimal expansion of $\{10x\}$, which is simply the second digit x_2 of the decimal expansion of x. Continuing in this way, we see that the symbolic sequence derived from x is none other than the sequence of digits from its decimal expansion.

Symbolic dynamics began when Jacques Hadamard [Had] applied this idea in 1898 to more complicated systems called geodesic flows on surfaces of negative curvature. The main point of his work is that there is a simple description of the possible sequences that can arise this way. He showed that there is a finite set of forbidden pairs of symbols, and that the possible sequences are exactly those that do not contain any forbidden pair. This is an example of one of the fundamental objects of study in symbolic dynamics called a shift of finite type. Later discoveries of Morse, Hedlund, and others in the 1920's, 1930's, and 1940's showed that in many circumstances such a finite description of the dynamics is possible. These ideas led in the 1960's and 1970's to the development of powerful mathematical tools to investigate a class of extremely interesting mappings called hyperbolic diffeomorphisms. We will see in §6.5 some examples of this.

Another source of inspiration and questions in symbolic dynamics comes from data storage and transmission. As we discuss in §2.5, when information is stored as bits on a magnetic or optical disk, there are physical reasons why it is not stored verbatim. For example, the bits on the surface of a compact audio disk are written in a long sequence obeying the constraint that between successive 1's there are at least two 0's but no more than ten 0's. How can one efficiently transform arbitrary data (such as a computer program or a Beethoven symphony) into sequences that satisfy such kinds of constraints? What are the theoretical limits of this kind of transformation? We are again confronted with a space of sequences having a finite description, and we are asking questions about such spaces and ways to encode and decode data from one space (the space of arbitrary sequences) to another (the space of constrained sequences). One of the main results in this book, the Finite-State Coding Theorem of Chapter 5, tells us when such codes are possible, and gives us an algorithm for finding them. This has led to new codes and a deeper understanding of code constructions. In particular, it has yielded a new and useful technique for code construction called the state-splitting algorithm.

While symbolic dynamics has drawn heavily on other mathematical disciplines such as linear algebra for its tools, it has also contributed new tools for other areas. For example, some deep work of Boyle and Handelman in symbolic dynamics described in Chapter 11 has led to the complete solution of the problem of characterizing the possible sets of nonzero eigenvalues of real nonnegative matrices. This solved a variant of a problem



Preface to first edition

xv

which has been perplexing linear algebraists for decades.

This book is intended to serve as an introduction to symbolic dynamics, its basic ideas, techniques, problems, and spirit. We will focus on symbolic dynamical systems that use just a finite number of symbols, on sequences of such symbols that are infinite in both directions (unlike the decimal expansion above), and spaces of sequences that have finite memory properties. Our aim is to study coding problems for such systems. In particular, we concentrate on the following:

- Find complete sets of necessary and sufficient conditions for the existence of various types of codes from one symbolic dynamical system to another.
- Completely determine the values of certain properties, invariant under a natural notion of conjugacy, such as entropy, numbers of periodic sequences, and so on.
- Explore interactions with information and coding theory, linear algebra, and general dynamical systems.

On the other hand, there are many important topics within symbolic dynamics that we do not treat in any depth in this book. Many of these topics are briefly discussed in Chapter 13.

The book is organized as follows. We start in Chapter 1 with the basic notions in symbolic dynamical systems, such as full shifts, shift spaces, irreducibility, and sliding block codes. In Chapter 2 we focus on a special class of shift spaces called shifts of finite type, and in Chapter 3 we study a generalization of these called sofic shifts. Entropy, an important numerical invariant, is treated in Chapter 4, which also includes an introduction to the fundamental Perron-Frobenius theory of nonnegative matrices. Building on the material in the first four chapters, we develop in Chapter 5 the state-splitting algorithm for code construction and prove the Finite-State Coding Theorem. Taken together, the first five chapters form a self-contained introduction to symbolic dynamics and its applications to practical coding problems.

Starting with Chapter 6 we switch gears. This chapter provides some background for the general theory of dynamical systems and the place occupied by symbolic dynamics. We show how certain combinatorial ideas like shift spaces and sliding block codes studied in earlier chapters are natural expressions of fundamental mathematical notions such as compactness and continuity. There is a natural notion of two symbolic systems being "the same," or conjugate, and Chapter 7 deals with the question of when two shifts of finite type are conjugate. This is followed in Chapters 8 and 9 with the classification of shifts of finite type and sofic shifts using weaker notions, namely finite equivalence and almost conjugacy. Chapters 7, 8, and 9 treat problems of coding between systems with equal entropy. In Chapter 10 we treat the case of unequal entropy and, in particular, deter-



xvi

Preface to first edition

mine when one shift of finite type can be embedded in another. The set of numbers which can occur as entropies of shifts of finite type is determined in Chapter 11. Also in that chapter we draw on the results of Chapter 10 to prove a partial result towards characterizing the zeta functions of shifts of finite type (the zeta function is a function which determines the numbers of periodic sequences of all periods). Some of these results in turn are used in Chapter 12 to classify shifts of finite type and sofic shifts up to a notion that sits in between conjugacy and almost conjugacy. The main result of Chapter 12 may be regarded as a generalization of the Finite-State Coding Theorem, and tells us when one sofic shift can be encoded into another in a natural way. Chapter 13 contains a survey of more advanced results and recent literature on the subject. This is a starting point for students wishing to become involved in areas of current research.

The mathematical prerequisites for this book are relatively modest. A reader who has mastered undergraduate courses in linear algebra, abstract algebra, and calculus should be well prepared. This includes upper-division undergraduate mathematics majors, mathematics graduate students, and graduate students studying information and storage systems in computer science and electrical engineering. We have tried to start from scratch with basic material that requires little background. Thus in the first five chapters no more than basic linear algebra and one-variable calculus is needed: matrices, linear transformations, similarity, eigenvalues, eigenvectors, and notions of convergence and continuity (there are also a few references to basic abstract algebra, but these are not crucial). Starting with Chapter 6, the required level of mathematical sophistication increases. We hope that the foundation laid in the first five chapters will inspire students to continue reading. Here the reader should know something more of linear algebra, including the Jordan canonical form, as well as some abstract algebra including groups, rings, fields, and ideals, and some basic notions from the topology of Euclidean space such as open set, closed set, and compact set, which are quickly reviewed.

The book contains over 500 exercises, ranging from simple checks of understanding to much more difficult problems and projects. Those that seem harder have been marked with a *. There is a detailed index as well as a notation index.

Theorems, propositions, lemmas, and other items are numbered consecutively within sections using three parts: k.m.n refers to item n in section m of Chapter k. For instance, Proposition 5.2.3 refers to the third item in Section 2 of Chapter 5. Equations are numbered separately with the same three-part system, but with a different style, so that (3-2-4) refers to the fourth equation in Section 2 of Chapter 3.

We use the standard notations of \mathbb{Z} , \mathbb{Q} , \mathbb{R} , and \mathbb{C} for the sets of integers, rationals, reals, and complexes, and use \mathbb{Z}^+ , \mathbb{Q}^+ , and \mathbb{R}^+ to denote the nonnegative elements of the corresponding sets. All matrices are assumed



Preface to first edition

xvii

to be square unless stated (or context demands) otherwise.

The authors have taught courses based on preliminary versions of this material at the University of Washington, Stanford University and the University of California at Berkeley. They have also given lecture series at the Mathematical Sciences Research Center in Berkeley and the IBM Almaden Research Center. Based on these experiences, here are some suggestions on how to organize the material in this book into courses of different lengths. Chapters 1–5 would constitute a solid one-quarter course introducing symbolic dynamics from scratch and concluding with applications to coding. Chapters 1–7 could form a one-semester course that also includes more theoretical material and finishes with a detailed discussion of the conjugacy problem. The entire book can be covered in a one-year course that would bring students to the frontiers of current research. Of course, the material can be covered more quickly for students with more than the minimum required background.

This book has benefited enormously from suggestions, comments, and corrections made by many people (students, faculty and industrial researchers) on preliminary versions of this manuscript, in lectures on these versions, and in answers to our queries. These include Paul Algoet. Jonathan Ashley, Leslie Badoian, Mignon Belongie, Mike Boyle, Karen Brucks, Elise Cawley, Phil Chou, Wu Chou, Ethan Coven, Jim Fitzpatrick, Dave Forney, Lisa Goldberg, Michael Gormish, Michael Hollander, Danrun Huang, Natasha Jonoska, Sampath Kannan, Bruce Kitchens, Wolfgang Krieger, David Larsen, Nelson Markley, Lee Neuwirth, Kyewon Koh Park, Karl Petersen, Ronny Roth, Paul Siegel, Sylvia Silberger, N. T. Sindhushayana, Serge Troubetzkoy, Paul Trow, Selim Tuncel, Jeff Von Limbach, Jack Wagoner, Peter Walters, Barak Weiss, Susan Williams, and Amy Wilkinson. We are happy to thank Roy Adler, who first suggested the possibility of a textbook on symbolic dynamics, especially one that requires only a minimum amount of mathematical background. Adler's 1984 set of unpublished lecture notes from Brown University and his notes on Markov partitions were valuable sources. We are also grateful to Elza Erkip, Amos Lapidoth, and Erik Ordentlich, who earned their just desserts by working so many of the exercises in the first five chapters. Special thanks are due to Zhe-xian Wan for very detailed and helpful comments on our manuscript and to Brian Hopkins for clarifying a multitude of murky passages and for eliminating an uncountable number of needless commas. This book is typeset using the LAMS-TEX extensions to TEX, and we thank Michael Spivak for his help with this. Finally, we wish to express our gratitude to the National Science Foundation (grants DMS-9004252 and DMS-9303240), the University of Washington, and the IBM Corporation for their support of this work.

Current information regarding this book, including selected corrections and suggestions from readers, may be found using the World Wide Web at the address



xviii

Preface to first edition

https://sites.math.washington.edu/SymbolicDynamics

This address can also be used to submit items for inclusion.

The first author is profoundly grateful to John Whittier Treat for many things, both spoken and unspoken, over many years.

The second author would like to thank Yvonne, Nathaniel, Mom and Dad for a life filled with love and support, without which this book could not have been written.

Notes on the Second Printing

The Second Printing (1999) has provided us the opportunity to fix some minor errors. In addition, we have made a few changes to reflect the recent solution by Kim and Roush [KimR12] of the Shift Equivalence Problem described in §7.3.

A complete list of changes made for the Second Printing is available at the Web site displayed above. We will continue to list further corrections at this Web site.



PREFACE TO SECOND EDITION

The twenty five years since our book was written have seen enormous advances in areas related to symbolic dynamics, with long-standing problems solved and completely new research directions undergoing vigorous development. We wanted to preserve the introductory character of the original book, while providing an extended account of many of these new topics in a separate Addendum. Therefore we restricted changes in our original book to clarifications and correcting minor errors, together with noting the solutions to the Spectral Conjecture and the Road Problem.

The Addendum surveys topics that appear to us to be important and interesting. These include the characterization of the entropies of higher dimensional shifts of finite type, the pivotal use of ideas from computation theory and logic to attack dynamical problems in higher dimensional shifts, and the extension of entropy theory to actions of sofic groups. So many new and interesting developments are happening that we could not possibly be comprehensive. But we have provided an extensive separate Addendum Bibliography that we hope will guide the reader to find out more about something that catches their eye.

We are very grateful for the help we have received from Sebastián Barbieri, Lewis Bowen, Raimundo Briceño, Justin Cai, Tullio Ceccherini-Silberstein, Nishant Chandgotia, Michel Coornaert, Madeline Doering, Søren Eilers, Rafael Frongillo, Ricardo Gomez, Ben Hayes, David Kerr, Bruce Kitchens, Bryna Kra, Hanfeng Li, Tom Meyerovitch, Ronnie Pavlov, Ian Putnam, Brendan Seward, and Marcelo Sobottka.

We are especially indebted to Mike Boyle, whose list of open problems in symbolic dynamics played an essential part in structuring the Addendum, and whose encyclopedic knowledge guided us at many points to provide perspectives we hope readers will find valuable.

We are also extremely grateful to Luisa Borsato and Sophie MacDonald for their careful reading of our original text and making many excellent suggestions and corrections.

Reconstructing this book from the original TEX files (from an incomplete backup) using the \LaTeX -TEX package was only possible with the help of Ian Allison and Will Kranz.

Finally, it is a great pleasure to thank Roger Astley for his initiative in proposing this second edition, Clare Dennison for overseeing production of the book, and Irene Pizzie for her meticulous copyediting.

 $\begin{array}{c} \text{Douglas Lind, } \textit{Seattle} \\ \text{Brian Marcus, } \textit{Vancouver} \\ \text{September, 2020} \end{array}$