**Classics in Mathematics** 

Richard S. Ellis Entropy, Large Deviations, and Statistical Mechanics

**Richard S. Ellis** 

# Entropy, Large Deviations, and Statistical Mechanics

Reprint of the 1985 Edition



Richard S. Ellis Department of Mathematics and Statistics University of Massachusetts Amherst, Massachusetts 01003 U.S.A.

Originally published as Vol. 271 in the series Grundlehren der mathematischen Wissenschaften

Mathematics Subject Classification (2000): 82A05, 60K35

Library of Congress Control Number: 2005934786

ISSN 1431-0821 ISBN-10 3-540-29059-1 Springer Berlin Heidelberg New York ISBN-13 978-3-540-29059-9 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2006 Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Production: LE-T<sub>E</sub>X Jelonek, Schmidt & Vöckler GbR, Leipzig Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper 41/3142/YL - 5 4 3 2 1 0

## Grundlehren der mathematischen Wissenschaften 271

A Series of Comprehensive Studies in Mathematics

Editors

M. Artin S. S. Chern J. M. Fröhlich E. Heinz H. Hironaka F. Hirzebruch L. Hörmander S. Mac Lane W. Magnus C. C. Moore J. K. Moser M. Nagata W. Schmidt D. S. Scott Ya. G. Sinai J. Tits B. L. van der Waerden M. Waldschmidt S. Watanabe

Managing Editors

M. Berger B. Eckmann S. R. S. Varadhan

#### Grundlehren der mathematischen Wissenschaften

A Series of Comprehensive Studies in Mathematics

A Selection

- 190. Faith: Algebra: Rings, Modules, and Categories I
- 191. Faith: Algebra II, Ring Theory
- 192. Mal'cev: Algebraic Systems
- 193. Pólya/Szegö: Problems and Theorems in Analysis I
- 194. Igusa: Theta Functions
- 195. Berberian: Baer\*-Rings
- 196. Athreya/Ney: Branching Processes
- 197. Benz: Vorlesungen über Geometrie der Algebren
- 198. Gaal: Linear Analysis and Representation Theory
- 199. Nitsche: Vorlesungen über Minimalflächen
- 200. Dold: Lectures on Algebraic Topology
- 201. Beck: Continuous Flows in the Plane
- 202. Schmetterer: Introduction to Mathematical Statistics
- 203. Schoeneberg: Elliptic Modular Functions
- 204. Popov: Hyperstability of Control Systems
- 205. Nikol'skii: Approximation of Functions of Several Variables and Imbedding Theorems
- 206. André: Homologie des Algèbres Commutatives
- 207. Donoghue: Monotone Matrix Functions and Analytic Continuation
- 208. Lacey: The Isometric Theory of Classical Banach Spaces
- 209. Ringel: Map Color Theorem
- 210. Gihman/Skorohod: The Theory of Stochastic Processes I
- 211. Comfort/Negrepontis: The Theory of Ultrafilters
- 212. Switzer: Algebraic Topology-Homotopy and Homology
- 213. Shafarevich: Basic Algebraic Geometry
- 214. van der Waerden: Group Theory and Quantum Mechanics
- 215. Schaefer: Banach Lattices and Positive Operators
- 216. Pólya/Szegö: Problems and Theorems in Analysis II
- 217. Stenström: Rings of Quotients
- 218. Gihman/Skorohod: The Theory of Stochastic Process II
- 219. Duvant/Lions: Inequalities in Mechanics and Physics
- 220. Kirillov: Elements of the Theory of Representations
- 221. Mumford: Algebraic Geometry I: Complex Projective Varieties
- 222. Lang: Introduction to Modular Forms
- 223. Bergh/Löfström: Interpolation Spaces. An Introduction
- 224. Gilbarg/Trudinger: Elliptic Partial Differential Equations of Second Order
- 225. Schütte: Proof Theory
- 226. Karoubi: K-Theory, An Introduction
- 227. Grauert/Remmert: Theorie der Steinschen Räume
- 228. Segal/Kunze: Integrals and Operators
- 229. Hasse: Number Theory
- 230. Klingenberg: Lectures on Closed Geodesics
- 231. Lang: Elliptic Curves: Diophantine Analysis
- 232. Gihman/Skorohod: The Theory of Stochastic Processes III
- 233. Stroock/Varadhan: Multi-dimensional Diffusion Processes
- 234. Aigner: Combinatorial Theory

Richard S. Ellis

# Entropy, Large Deviations, and Statistical Mechanics



Springer-Verlag New York Berlin Heidelberg Tokyo Richard S. Ellis Department of Mathematics and Statistics University of Massachusetts Amherst, Massachusetts 01003 U.S.A.

#### AMS Subject Classifications: 82A05, 60K35

Library of Congress Cataloging in Publication Data Ellis, Richard S. (Richard Steven) Entropy, large deviations, and statistical mechanics. (Grundlehren der mathematischen Wissenschaften; 271) Bibliography: p. Includes index. 1. Statistical mechanics. 2. Entropy. 3. Large deviations. I. Title. II. Series. QC174.8.E45 1985 530.1'3 84-13891

With 8 Illustrations

© 1985 by Springer-Verlag New York Inc. All rights reserved. No part of this book may be translated or reproduced in any form without written permission from Springer-Verlag, 175 Fifth Avenue, New York, New York 10010, U.S.A.

Typeset by Asco Trade Typesetting Ltd., Hong Kong. Printed and bound by R. R. Donnelley and Sons, Harrisonburg, Virginia. Printed in the United States of America.

987654321

ISBN 0-387-96052-X Springer-Verlag New York Berlin Heidelberg Tokyo ISBN 3-540-96052-X Springer-Verlag Berlin Heidelberg New York Tokyo For Alison, Melissa, and Michael my wife and children who make it all worthwhile

### Preface

This book has two main topics: large deviations and equilibrium statistical mechanics. I hope to convince the reader that these topics have many points of contact and that in being treated together, they enrich each other. Entropy, in its various guises, is their common core.

The large deviation theory which is developed in this book focuses upon convergence properties of certain stochastic systems. An elementary example is the weak law of large numbers. For each positive  $\varepsilon$ ,  $P\{|S_n/n| \ge \varepsilon\}$  converges to zero as  $n \to \infty$ , where  $S_n$  is the *n*th partial sum of independent identically distributed random variables with zero mean. Large deviation theory shows that if the random variables are exponentially bounded, then the probabilities converge to zero exponentially fast as  $n \to \infty$ . The exponential decay allows one to prove the stronger property of almost sure convergence  $(S_n/n \rightarrow 0 \text{ a.s.})$ . This example will be generalized extensively in the book. We will treat a large class of stochastic systems which involve both independent and dependent random variables and which have the following features: probabilities converge to zero exponentially fast as the size of the system increases; the exponential decay leads to strong convergence properties of the system. The most fascinating aspect of the theory is that the exponential decay rates are computable in terms of entropy functions. This identification between entropy and decay rates of large deviation probabilities enhances the theory significantly.

Entropy functions have their roots in statistical mechanics. They originated in the work of L. Boltzmann, who in the 1870's studied the relation between entropy and probability in physical systems. Thus statistical mechanics has a strong historical connection with large deviation theory. It also provides a natural context in which the theory can be applied. Applications of large deviations to models in equilibrium statistical mechanics are presented in Chapters III–V. These applications illustrate convincingly the power of the theory.

Equilibrium statistical mechanics is an exciting area of mathematical physics but one which remains inaccessible to many mathematicians. Some texts on the subject provide an introduction to the physics but do not develop the mathematics in much detail or with great rigor. Other texts treat mathematical problems in statistical mechanics with complete rigor but assume an extensive background in the physics. The uninitiated reader has difficulty understanding how concepts like ensemble, free energy, or entropy connect up with more familiar concepts in mathematics. My approach in this book is to emphasize strongly the connections between statistical mechanics on the one hand and probability and large deviations on the other. I hope that in so doing, I have succeeded in providing a readable treatment of statistical mechanics which is accessible to a general mathematical audience. My large deviation approach to statistical mechanics was inspired in part by the article of O. E. Lanford (1973).

In recent years, the scope of large deviations has been greatly expanded by M. D. Donsker and S. R. S. Varadhan. This book contains an introduction to their theory. I illustrate the main features in the context of independent identically distributed random vectors taking values in  $\mathbb{R}^d$ . I also present my own large deviation results, which are particularly suited for applications to statistical mechanics. Since readability rather than completeness has been my goal, the large deviation theorems are not stated in the greatest generality.

There are two parts to the book, Part I consisting of Chapters I–V. Chapter I introduces large deviations by means of elementary examples involving combinatorics and Stirling's formula. Chapter II presents the Donsker–Varadhan theory as well as my own large deviation results. The proofs of the theorems in this chapter are detailed and are postponed until Part II. Postponing proofs allows the reader to reach, as soon as possible, interesting applications of large deviation analysis of a discrete gas model. Chapters IV–V discuss the Ising model of ferromagnetism and related spin systems. The emphasis in these two chapters is upon properties of Gibbs states. While large deviation theory provides a terminology and a set of results that are useful for treating Gibbs states, the book also develops other tools that are needed. These include convexity and moment inequalities.

Part II consists of Chapters VI–IX. Chapter VI is a summary of the theory of convex functions on  $\mathbb{R}^d$ . Chapters VII–IX prove the large deviation results stated in Chapter II without proof. The prerequisite for these chapters is a good working knowledge of probability and measure theory. The essential definitions and theorems in probability are listed in Appendix A. The appendix is intended to be a review or an outline for study rather than a detailed exposition.

This book can be used as a text. It contains over 100 problems, many of which have hints. Chapters I and II and VI–IX are a self-contained treatment of large deviations and convex functions. Readers primarily interested in spin systems can concentrate upon Chapters IV and V and refer to the statements and proofs of large deviation results as needed. Those portions of Chapters IV and V which do not rely on large deviations are self-contained. Chapters IV and V can be completely understood without reading Chapter III.

This book contains new results and new proofs of known theorems. These include the following: exponential convergence properties of Gibbs states

[Theorems IV.5.5, IV.6.6, and V.6.1]; a large deviation proof of the Gibbs variational formula [Theorem IV.7.3(a)]; a proof of the central limit theorem for spin systems [Theorem V.7.2(a)]; a level-3 large deviation theorem for i.i.d. random variables with a finite state space [Theorem IX.1.1]; a level-3 large deviation theorem for Markov chains with a finite state space [Problems IX.6.10–IX.6.15]; the solution of the Gibbs variational formula for finite-range interactions on  $\mathbb{Z}$  via large deviations [Appendix C.6]. Many of the large deviation results and applications in the book depend upon my large deviation theorem, Theorem II.6.1. The proof of the level-3 theorem in Chapter IX was inspired by statistical mechanics [see Appendix C.6] and information theory.

I have had the good fortune of interacting with a number of special people. Todd Baker edited the manuscript with creativity and care. The book benefited greatly from his involvement. Peg Bombardier was my superb typist. She was always cheerful and patient, despite the numerous revisions, and was a pleasure to work with. Alan Sokal read portions of the manuscript and was a big help with the statistical mechanics. I owe a special debt of gratitude to Srinivasa Varadhan. He answered my many questions about large deviations patiently and with insight and showed a strong interest in the book. The encouragement of my family and friends was greatly appreciated. Above all, I thank my wife Alison. Her love is a blessing.

I am grateful to Alejandro de Acosta, Hans-Otto Georgii, Joseph Horowitz, Jonathan Machta, Charles Newman, and R. Tyrrell Rockafellar for reading portions of the manuscript and suggesting improvements. I am also indebted to the many other people, too numerous to mention by name, with whom I have consulted. While writing the book, I received support from the University of Massachusetts, the National Science Foundation, and the Lady Davis Fellowship Trust. Their support is gratefully acknowledged.

Richard S. Ellis

#### Comments on the Use of This Book

At the end of each chapter, there is a Notes section, followed by a Problems section. References to the Notes are indicated by superscripted integers; e.g., entropy<sup>4</sup> refers to Note 4. Near the end of the book, there is a list of frequently used symbols.

The main large deviation theorems are stated in Chapter II and are proved in Chapters VI–IX. Readers interested primarily in large deviations may read Chapters I, II, VI–IX while those interested primarily in statistical mechanics may read Chapters I–V.

### Contents

Preface	vii
Comments on the Use of This Book	x
PART I: LARGE DEVIATIONS AND STATISTICAL MECHAN	VICS
Chapter I. Introduction to Large Deviations	3
I.1. Overview	3
I.2. Large Deviations for I.I.D. Random Variables with a Finite State	-
Space	9
I.3. Levels-1 and 2 for Coin Tossing	11
I.4. Levels-1 and 2 for I.I.D. Random Variables with a Finite State	
Space	13
I.5. Level-3: Empirical Pair Measure	17
I.6. Level-3: Empirical Process	21
I.7. Notes	26
I.8. Problems	27
Chapter II. Large Deviation Property and Asymptotics of Integrals	30
II.1. Introduction	30
II.2. Levels-1, 2, and 3 Large Deviations for I.I.D. Random Vectors	30
II.3. The Definition of Large Deviation Property	33
II.4. Statement of Large Deviation Properties for Levels-1, 2, and 3	37
II.5. Contraction Principles	42
II.6. Large Deviation Property for Random Vectors and Exponential	
Convergence	46
II.7. Varadhan's Theorem on the Asymptotics of Integrals	50
II.8. Notes	55
II.9. Problems	56
Chapter III. Large Deviations and the Discrete Ideal Gas	59
III.1. Introduction	59
III.2. Physics Prelude: Thermodynamics	60
III.3. The Discrete Ideal Gas and the Microcanonical Ensemble	64

III.4.	Thermodynamic Limit, Exponential Convergence, and	
	Equilibrium Values	66
III.5.	The Maxwell–Boltzmann Distribution and Temperature	74
III.6.	The Canonical Ensemble and Its Equivalence with the	
	Microcanonical Ensemble	75
III.7.	A Derivation of a Thermodynamic Equation	78
III.8.	The Gibbs Variational Formula and Principle	79
III.9.	Notes	84
III.10.	Problems	86
Chapter	: IV. Ferromagnetic Models on $\mathbb{Z}$	88
IV.1.	Introduction	88
IV.2.	An Overview of Ferromagnetic Models	88
IV.3.	Finite-Volume Gibbs States on $\mathbb{Z}$	94
IV.4.	Spontaneous Magnetization for the Curie–Weiss Model	98
IV.5.	Spontaneous Magnetization for General Ferromagnets on $\mathbb{Z}$	103
IV.6.	Infinite-Volume Gibbs States and Phase Transitions	109
IV.7.	The Gibbs Variational Formula and Principle	125
IV.8.	Notes	131
IV.9.	Problems	134
Chapter	V. Magnetic Models on $\mathbb{Z}^{D}$ and on the Circle	138
V.1.	Introduction	138
V.2.	Finite-Volume Gibbs States on $\mathbb{Z}^{D}$ , $D \geq 1$	140
<b>V</b> .3.	Moment Inequalities	142
V.4.	Properties of the Magnetization and the Gibbs Free Energy	146
V.5.	Spontaneous Magnetization on $\mathbb{Z}^{D}$ , $D \geq 2$ , Via the Peierls Argument	153
V.6.	Infinite-Volume Gibbs States and Phase Transitions	157
<b>V</b> .7.	Infinite-Volume Gibbs States and the Central Limit Theorem	162
<b>V.8</b> .	Critical Phenomena and the Breakdown of the Central Limit	
	Theorem	170
V.9.	Three Faces of the Curie–Weiss Model	179
V.10.	The Circle Model and Random Waves	190
V.11.	A Postscript on Magnetic Models	198
V.12.	Notes	199
V.13.	Problems	- 203

## PART II: CONVEXITY AND PROOFS OF LARGE DEVIATION THEOREMS

Chapter	VI. Convex Functions and the Legendre-Fenchel Transform	211
VI.1.	Introduction	211
VI.2.	Basic Definitions	211
VI.3.	Properties of Convex Functions	213

xii

<ul> <li>VI.4. A One-Dimensional Example of the Legendre–Fenchel Transform</li> <li>VI.5. The Legendre–Fenchel Transform for Convex Functions on ℝ<sup>d</sup></li> <li>VI.6. Notes</li> <li>VI.7. Problems</li> </ul>	218 220 225 226
Chapter VII. Large Deviations for Random Vectors	229
VII.1. Statement of Results	229
VII.2. Properties of $I_{\mathcal{W}}$	231
VII.3. Proof of the Large Deviation Bounds for $d = 1$	232
VII.4. Proof of the Large Deviation Bounds for $d \ge 1$	235
VII.5. Level-1 Large Deviations for I.I.D. Random Vectors	238
VII.6. Exponential Convergence and Proof of Theorem II.6.3	242
VII.7. Notes	243
VII.8. Problems	245
Chapter VIII. Level-2 Large Deviations for I.I.D. Random Vectors VIII.1. Introduction VIII.2. The Level-2 Large Deviation Theorem VIII.3. The Contraction Principle Relating Levels-1 and 2 ( $d = 1$ ) VIII.4. The Contraction Principle Relating Levels-1 and 2 ( $d \ge 2$ ) VIII.5. Notes VIII.6. Problems	250 250 253 258 264 265
Chapter IX. Level-3 Large Deviations for I.I.D. Random Vectors	269
IX.1. Statement of Results	269
IX.2. Properties of the Level-3 Entropy Function	270
IX.3. Contraction Principles	276
IX.4. Proof of the Level-3 Large Deviation Bounds	279
IX.5. Notes	287
IX.6. Problems	288
APPENDICES	
APPENDICES	

Append	lix A: Probability	295
A.1.	Introduction	295
A.2.	Measurability	295
A.3.	Product Spaces	296
A.4.	Probability Measures and Expectation	297
A.5.	Convergence of Random Vectors	299
A.6.	Conditional Expectation, Conditional Probability, and Regular	
	Conditional Distribution	299
A.7.	The Kolmogorov Existence Theorem	301

xiii

A.8. A.9. A 10	Weak Convergence of Probability Measures on a Metric Space The Space $\mathcal{M}_s((\mathbb{R}^d)^{\mathbb{Z}})$ and the Ergodic Theorem <i>n</i> -Dependent Markov Chains	303 306 311
A.11.	Probability Measures on the Space $\{1, -1\}^{\mathbb{Z}^D}$	312
Append	dix B: Proofs of Two Theorems in Section II.7	319
B.1. B.2.	Proof of Theorem II.7.1 Proof of Theorem II.7.2	319 321
Appen	dix C: Equivalent Notions of Infinite-Volume Measures for Spin Systems	323
C.1.	Introduction	323
C.2.	Two-Body Interactions and Infinite-Volume Gibbs States	323
C.3.	Many-Body Interactions and Infinite-Volume Gibbs States	325
C.4.	DLR States	326
C.5.	I ne Globs variational Formula and Principle	321
C.0.	Interactions on $\mathbb{Z}$	330
Appen	dix D: Existence of the Specific Gibbs Free Energy	332
D.1.	Existence Along Hypercubes	332 334
10.2.		
List of	Frequently Used Symbols	335
Referen	nces	338
Author	Index	353
Subjec	t Index	359

Part I

## Large Deviations and Statistical Mechanics