

FIELDS INSTITUTE MONOGRAPHS

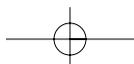
THE FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCES

Large Deviations

Frank den Hollander



American Mathematical Society
Providence, Rhode Island



PREFACE

Large deviation theory is a part of probability theory that deals with the description of events where a sum of random variables deviates from its mean by more than a “normal” amount, i.e., beyond what is described by the central limit theorem. A precise calculation of the probabilities of such events turns out to be crucial for the study of integrals of exponential functionals of sums of random variables, which come up in a variety of different contexts. Large deviation theory finds application in probability theory, statistics, operations research, ergodic theory, information theory, statistical physics, financial mathematics, and the list goes on.

These lecture notes are an *introduction* to large deviations. Part A (Chapters I–V) describes *theory*, Part B (Chapters VI–X) describes *applications*. A glance at the table of contents shows what topics are covered. I have put much effort into conveying the main ideas without putting too much emphasis on technicalities. Most of the theory is driven by a few “key principles” and once these are understood the rest of the journey is safe sailing, give or take a storm or two. This is not to say that it is easy to grasp the full theoretical panorama. But the reader’s patience will be rewarded when the ship enters the harbor of the applications.

Chapters I and II contain the basic large deviation theorems for i.i.d. random variables. Here the goal is to make the reader acquainted with the type of statements that are typical for the theory and to obtain results via explicit calculation of the rate function. Chapter III presents general definitions and theorems in a more abstract context. Here the goal is to expose the unified scheme that gives large deviation theory its overall structure and that can be made to work in many concrete cases. Chapter IV looks at large deviation theorems for Markov chains and explains how these can be obtained from the i.i.d. case via a change-of-measure argument. Chapter V considers random sequences with moderate dependence and shows how many of the results in Chapters I, II and IV can be put under a single heading, which in some sense closes the circle.

Chapters VI–X describe a selection of applications: statistical hypothesis testing; random walk in random environment; heat conduction with random sources and sinks; polymer chains; interacting diffusions. Here the theory comes to life and the reader gets to see the full impact of the results derived earlier. Except for the application in statistical hypothesis testing, which is put in mainly for didactical reasons, all applications are recent. Naturally their choice reflects my personal taste and involvement, since they circle around statistical physics and random media. But I think they offer a good sample of what large deviation theory is able to do in various different contexts. Each application is self-contained and tells a small story.

Many questions that come up during the exposition are posed as exercises to the reader. The solutions to these exercises are given in the Appendix. At the end I have included a list of frequently used words and symbols, with the number of the section where they appear first. This will help the reader to connect the different chapters.

Large deviation theory is a mixture of probability theory, convex analysis, variational calculus and set topology. As such it is mathematically both challenging and captivating. Even so, it is not obvious how to do justice to a vast area like large deviation theory in one hundred pages or so, especially when the goal is to cover both theory and applications. To focus ideas, I have restricted most of the exposition to random sequences, i.e., discrete-time random processes. This is a severe restriction indeed, but it makes the presentation much more user-friendly. The reader can expand his or her skills by turning to the monographs that are listed as references. Here a wealth of refinements and embellishments can be found, as well as beautiful and deep large deviation results for Brownian motion, random dynamical systems, Gibbs measures, interacting particle systems, Brownian motion among random obstacles, etc. These monographs also contain an extensive historical overview of the area.

The material presented here was taught as a graduate course at the Fields Institute for Research in Mathematical Sciences in Toronto, Canada, in the Fall of 1998, as part of the 1998-99 program on “Probability and its Applications”. I am grateful to the staff of the Fields Institute for the hospitality I enjoyed as a visitor. I am grateful to the following colleagues for comments during the course: Rami Atar, Siva Athreya, Marek Biskup, Jürgen Gärtner, Takashi Hara, Remco van der Hofstad, Min Kang, Neal Madras, Anders Martin-Löf, George O’Brien, David Rolls, Tom Salisbury, Gordon Slade, Dean Slonowsky, Jan Swart and Stas Volkov.

I am indebted to Siva Athreya, Nina Gantert, Jürgen Gärtner, Andreas Greven, Remco van der Hofstad, Wolfgang König and George O’Brien for reading a draft of the manuscript and for suggesting improvements.

Special thanks go to Marek Biskup, who helped me to prepare the manuscript, both in terms of content and exposition, and who worked hard on the layout and on the figures. Marek was a constant companion in the enterprise of bringing the job to a good end.

Frank den Hollander
Nijmegen, October 1999

Contents

PREFACE	iii
Part A. THEORY	1
Chapter I. LARGE DEVIATIONS FOR I.I.D. SEQUENCES: PART 1	3
I.1. Introduction	3
I.2. An example: coin tossing	4
I.3. Cramér's Theorem for the empirical average	5
I.4. Comments	8
I.5. A toy application of Cramér	10
Chapter II. LARGE DEVIATIONS FOR I.I.D. SEQUENCES: PART 2	13
II.1. Sanov's Theorem for the empirical measure	13
II.2. The pair empirical measure	15
II.3. A toy application of Sanov for pairs	18
II.4. A contraction principle	19
II.5. The empirical process	20
II.6. Comments	25
II.7. Extension to countable state space	26
Chapter III. GENERAL THEORY	29
III.1. The large deviation principle (LDP)	29
III.2. Comments	31
III.3. Varadhan's Lemma	32
III.4. The LDP for integrals of exponential functionals	34
III.5. The Contraction Principle	35
III.6. The weak LDP	35
III.7. Convexity	36
III.8. Relation to earlier results	41
Chapter IV. LARGE DEVIATIONS FOR MARKOV SEQUENCES	43
IV.1. Radon-Nikodym formula	43
IV.2. The LDP for discrete-time Markov chains	44
IV.3. Comments	46
IV.4. The LDP for continuous-time Markov chains	47
Chapter V. LARGE DEVIATIONS FOR DEPENDENT SEQUENCES	53
V.1. Preliminaries	53
V.2. The Gärtner-Ellis Theorem	54
V.3. Comments	58
V.4. Relation to earlier results	60

V.5. Conclusion	61
Part B. APPLICATIONS	63
Chapter VI. STATISTICAL HYPOTHESIS TESTING	65
VI.1. The statistical problem	65
VI.2. Large deviation estimates on test optimality	66
Chapter VII. RANDOM WALK IN RANDOM ENVIRONMENT	69
VII.1. Random drifts	69
VII.2. The LDP for the speed	70
VII.3. The LDP for the hitting times	72
VII.4. From hitting times to speed	73
VII.5. Random continued fractions	74
VII.6. Analysis of the rate functions	75
VII.7. Comparison with homogeneous random walk	77
VII.8. Concluding remarks	79
Chapter VIII. HEAT CONDUCTION WITH RANDOM SOURCES AND SINKS	81
VIII.1. The parabolic Anderson model	81
VIII.2. Growth rate of the moments	83
VIII.3. Analysis of the variational problem	84
VIII.4. Correlation structure	85
VIII.5. Derivation of the growth rate	87
VIII.6. Transformation of the variational problem	89
VIII.7. Concluding remarks	91
Chapter IX. POLYMER CHAINS	93
IX.1. A polymer model: self-repellent random walk	93
IX.2. Linear speed	94
IX.3. The LDP for bridges	96
IX.4. Step 1: Adding drift	98
IX.5. Step 2: Markovian nature of the total local times	99
IX.6. Step 3: Key variational problem	100
IX.7. Step 4: Analysis of the rate function for bridges	102
IX.8. Step 5: Identification of the speed	104
IX.9. The LDP without the bridge condition	105
IX.10. Concluding remarks	108
Chapter X. INTERACTING DIFFUSIONS	111
X.1. The interaction Hamiltonian	111
X.2. Radon-Nikodym formula	112
X.3. The LDP for the double layer empirical measure	113
X.4. A simpler representation for the rate function	114
X.5. McKean-Vlasov equation	115
X.6. The Kuramoto model	117
X.7. Concluding remarks	120
Appendix: Solutions to the Exercises	121
Bibliography	135

Contents	vii
Index	139
Glossary of Symbols	141