



HAL
open science

Markov chains

Randal Douc, Eric Moulines, Pierre Priouret, Philippe Soulier

► **To cite this version:**

Randal Douc, Eric Moulines, Pierre Priouret, Philippe Soulier. Markov chains. Springer, pp.757, 2018, Operation research and financial engineering, Operation research and financial engineering, 978-3-319-97703-4. 10.1007/978-3-319-97704-1 . hal-02022651

HAL Id: hal-02022651

<https://hal.science/hal-02022651v1>

Submitted on 13 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Springer Series in Operations Research and Financial Engineering

Series editors

Thomas V. Mikosch
Sidney I. Resnick
Stephen M. Robinson

More information about this series at <http://www.springer.com/series/3182>

Randal Douc · Eric Moulines
Pierre Priouret · Philippe Soulier

Markov Chains

 Springer

Randal Douc
Département CITI
Telecom SudParis
Évry, France

Pierre Priouret
Université Pierre et Marie Curie
Paris, France

Eric Moulines
Centre de Mathématiques Appliquées
Ecole Polytechnique
Palaiseau, France

Philippe Soulier
Université Paris Nanterre
Nanterre, France

ISSN 1431-8598 ISSN 2197-1773 (electronic)
Springer Series in Operations Research and Financial Engineering
ISBN 978-3-319-97703-4 ISBN 978-3-319-97704-1 (eBook)
<https://doi.org/10.1007/978-3-319-97704-1>

Library of Congress Control Number: 2018950197

Mathematics Subject Classification (2010): 60J05, 60-02, 60B10, 60J10, 60J22, 60F05

© Springer Nature Switzerland AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, 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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Markov chains are a class of stochastic processes very commonly used to model random dynamical systems. Applications of Markov chains can be found in many fields, from statistical physics to financial time series. Examples of successful applications abound. Markov chains are routinely used in signal processing and control theory. Markov chains for storage and queueing models are at the heart of many operational research problems. Markov chain Monte Carlo methods and all their derivatives play an essential role in computational statistics and Bayesian inference.

The modern theory of discrete state-space Markov chains actually started in the 1930s with the work well ahead of its time of Doeblin (1938, 1940), and most of the theory (classification of states, existence of an invariant probability, rates of convergence to equilibrium, etc.) was already known by the end of the 1950s. Of course, there have been many specialized developments of discrete-state-space Markov chains since then, see for example Levin et al. (2009), but these developments are only taught in very specialized courses. Many books cover the classical theory of discrete-state-space Markov chains, from the most theoretical to the most practical. With few exceptions, they deal with almost the same concepts and differ only by the level of mathematical sophistication and the organization of the ideas.

This book deals with the theory of Markov chains on general state spaces. The foundations of general state-space Markov chains were laid in the 1940s, especially under the impulse of the Russian school (Yinnik, Yaglom, et al.). A summary of these early efforts can be found in Doob (1953). During the sixties and the seventies, some very significant results were obtained such as the extension of the notion of irreducibility, recurrence/transience classification, the existence of invariant measures, and limit theorems. The books by Orey (1971) and Foguel (1969) summarize these results.

Neveu (1972) brought many significant additions to the theory by introducing the taboo potential a function instead of a set. This approach is no longer widely used today in applied probability and will not be developed in this book (see, however, Chapter 4). The taboo potential approach was later expanded in the book

by Revuz (1975). This latter book contains much more and essentially summarizes all that was known in the mid seventies.

A breakthrough was achieved in the works of Nummelin (1978) and Athreya and Ney (1978), which introduce the notion of the split chain and embedded renewal process. These methods allow one to reduce the study to the case of Markov chains that possess an atom, that is, a set in which a regeneration occurs. The theory of such chains can be developed in complete analogy with discrete state space. The renewal approach leads to many important results, such as geometric ergodicity of recurrent Markov chains (Nummelin and Tweedie 1978; Nummelin and Tuominen 1982, 1983) and limit theorems (central limit theorems, law of iterated logarithms). This program was completed in the book Nummelin (1984), which contains a considerable number of results but is admittedly difficult to read.

This preface would be incomplete if we did not quote Meyn and Tweedie (1993b), referred to as the bible of Markov chains by P. Glynn in his prologue to the second edition of this book (Meyn and Tweedie 2009). Indeed, it must be acknowledged that this book has had a profound impact on the Markov chain community and on the authors. Three of us learned the theory of Markov chains from Meyn and Tweedie (1993b), which has therefore shaped and biased our understanding of this topic.

Meyn and Tweedie (1993b) quickly became a classic in applied probability and is praised by both theoretically inclined researchers and practitioners. This book offers a self-contained introduction to general state-space Markov chains, based on the split chain and embedded renewal techniques. The book recognizes the importance of Foster–Lyapunov drift criteria to assess recurrence or transience of a set and to obtain bounds for the return time or hitting time to a set. It also provides, for positive Markov chains, necessary and sufficient conditions for geometric convergence to stationarity.

The reason we thought it would be useful to write a new book is to survey some of the developments made during the 25 years that have elapsed since the publication of Meyn and Tweedie (1993b). To save space while remaining self-contained, this also implied presenting the classical theory of general state-space Markov chains in a more concise way, eliminating some developments that we thought are more peripheral.

Since the publication of Meyn and Tweedie (1993b), the field of Markov chains has remained very active. New applications have emerged such as Markov chain Monte Carlo (MCMC), which now plays a central role in computational statistics and applied probability. Theoretical development did not lag behind. Triggered by the advent of MCMC algorithms, the topic of quantitative bounds of convergence became a central issue. Much progress has been achieved in this field, using either coupling techniques or operator-theoretic methods. This is one of the main themes of several chapters of this book and still an active field of research. Meyn and Tweedie (1993b) deals only with geometric ergodicity and the associated Foster–Lyapunov drift conditions. Many works have been devoted to subgeometric rates of convergence to stationarity, following the pioneering paper of Tuominen and Tweedie (1994), which appeared shortly after the first version of Meyn and

Tweedie (1993b). These results were later sharpened in a series of works of Jarner and Roberts (2002) and Douc et al. (2004a), where a new drift condition was introduced. There has also been substantial activity on sample paths, limit theorems, and concentration inequalities. For example, Maxwell and Woodroffe (2000) and Rio (2017) obtained conditions for the central limit theorems for additive functions of Markov chains that are close to optimal.

Meyn and Tweedie (1993b) considered exclusively irreducible Markov chains and total variation convergence. There are, of course, many practically important situations in which the irreducibility assumption fails to hold, whereas it is still possible to prove the existence of a unique stationary probability and convergence to stationarity in distances weaker than the total variation. This quickly became an important field of research.

Of course, there are significant omissions in this book, which is already much longer than we initially thought it would be. We do not cover large deviations theory for additive functionals of Markov chains despite the recent advances made in this field in the work of Balaji and Meyn (2000) and Kontoyiannis and Meyn (2005). Similarly, significant progress has been made in the theory of moderate deviations for additive functionals of Markov chains in a series of Chen (1999), Guillin (2001), Djellout and Guillin (2001), and Chen and Guillin (2004). These efforts are not reported in this book. We do not address the theory of fluid limit introduced in Dai (1995) and later refined in Dai and Meyn (1995), Dai and Weiss (1996) and Fort et al. (2006), despite its importance in analyzing the stability of Markov chains and its success in analyzing storage systems (such as networks of queues). There are other significant omissions, and in many chapters we were obliged sometimes to make difficult decisions.

The book is divided into four parts. In Part I, we give the foundations of Markov chain theory. All the results presented in these chapters are very classical. There are two highlights in this part: Kac's construction of the invariant probability in Chapter 3 and the ergodic theorems in Chapter 5 (where we also present a short proof of Birkhoff's theorem).

In Part II, we present the core theory of irreducible Markov chains, which is a subset of Meyn and Tweedie (1993b). We use the regeneration approach to derive most results. Our presentation nevertheless differs from that of Meyn and Tweedie (1993b). We first focus on the theory of atomic chains in Chapter 6. We show that the atoms are either recurrent or transient, establish solidarity properties for atoms, and then discuss the existence of an invariant measure. In Chapter 7, we apply these results to discrete state spaces. We would like to stress that this book can be read without any prior knowledge of discrete-state-space Markov chains: all the results are established as a special case of atomic chains. In Chapter 8, we present the key elements of discrete time-renewal theory. We use the results obtained for discrete-state-space Markov chains to provide a proof of Blackwell and Kendall's theorems, which are central to discrete-time renewal theory. As a first application, we obtain a version of Harris's theorem for atomic Markov chains (based on the first-entrance last-exit decomposition) as well as geometric and polynomial rates of convergence to stationarity.

For Markov chains on general state spaces, the existence of an atom is more the exception than the rule. The splitting method consists in extending the state space to construct a Markov chain that contains the original Markov chain (as its first marginal) and has an atom. Such a construction requires that one have first defined small sets and petite sets, which are introduced in Chapter 9. We have adopted a definition of irreducibility that differs from the more common usage. This avoids the delicate theorem of Jain and Jamison (1967) (which is, however, proved in the appendix of this chapter for completeness but is not used) and allows us to define irreducibility on arbitrary state spaces (whereas the classical assumption requires the use of a countably generated σ -algebra). In Chapter 10, we discuss recurrence, Harris recurrence, and transience of general state-space Markov chains. In Chapter 11, we present the splitting construction and show how the results obtained in the atomic framework can be translated for general state-space Markov chains. The last chapter of this part, Chapter 12, deals with Markov chains on complete separable metric spaces. We introduce the notions of Feller, strong-Feller, and T -chains and show how the notions of small and petite sets can be related in such cases to compact sets. This is a very short presentation of the theory of Feller chains, which are treated in much greater detail in Meyn and Tweedie (1993b) and Borovkov (1998).

The first two parts of the book can be used as a text for a one-semester course, providing the essence of the theory of Markov chains but avoiding difficult technical developments. The mathematical prerequisites are a course in probability, stochastic processes, and measure theory at no deeper level than, for instance, Billingsley (1986) and Taylor (1997). All the measure-theoretic results that we use are recalled in the appendix with precise references. We also occasionally use some results from martingale theory (mainly the martingale convergence theorem), which are also recalled in the appendix. Familiarity with Williams (1991) or the first three chapters of Neveu (1975) is therefore highly recommended. We also occasionally need some topology and functional analysis results for which we mainly refer to the books Royden (1988) and Rudin (1987). Again, the results we use are recalled in the appendix.

Part III presents more advanced results for irreducible Markov chains. In Chapter 13, we complement the results that we obtained in Chapter 8 for atomic Markov chains. In particular, we cover subgeometric rates of convergence. The proofs presented in this chapter are partly original. In Chapter 14 we discuss the geometric regularity of a Markov chain and obtain the equivalence of geometric regularity with a Foster–Lyapunov drift condition. We use these results to establish geometric rates of convergence in Chapter 15. We also establish necessary and sufficient conditions for geometric ergodicity. These results are already reported in Meyn and Tweedie (2009). In Chapter 16, we discuss subgeometric regularity and obtain the equivalence of subgeometric regularity with a family of drift conditions. Most of the arguments are taken from Tuominen and Tweedie (1994). We then discuss the more practical subgeometric drift conditions proposed in Douc et al. (2004a), which are the counterpart of the Foster–Lyapunov conditions for geometric

regularity. In Chapter 17 we discuss the subgeometric rate of convergence to stationarity, using the splitting method.

In the last two chapters of this part, we reestablish the rates of convergence by two different types of methods that do not use the splitting technique.

In Chapter 18 we derive explicit geometric rates of convergence by means of operator-theoretic arguments and the fixed-point theorem. We introduce the uniform Doeblin condition and show that it is equivalent to uniform ergodicity, that is, convergence to the invariant distribution at the same geometric rate from every point of the state space. As a by-product, this result provides an alternative proof of the existence of an invariant measure for an irreducible recurrent kernel that does not use the splitting construction. We then prove nonuniform geometric rates of convergence by the operator method, using the ideas introduced in Hairer and Mattingly (2011).

In the last chapter of this part, Chapter 19, we discuss coupling methods that allow us to easily obtain quantitative convergence results as well as short and elegant proofs of several important results. We introduce different notions of coupling, starting almost from scratch: exact coupling, distributional coupling, and maximal coupling. This part owes much to the excellent treatises on coupling methods Lindvall and (1979) and Thorisson (2000), which of course cover much more than this chapter. We then show how exact coupling allows us to obtain explicit rates of convergence in the geometric and subgeometric cases. The use of coupling to obtain geometric rates was introduced in the pioneering work of Rosenthal (1995b) (some improvements were later supplied by Douc et al. (2004b)). We also illustrate the use of the exact coupling method to derive subgeometric rates of convergence; we follow here the work of Douc et al. (2006, 2007). Although the content of this part is more advanced, part of it can be used in a graduate course on Markov chains. The presentation of the operator-theoretic approach of Hairer and Mattingly (2011), which is both useful and simple, is of course a must. I also think it interesting to introduce the coupling methods, because they are both useful and elegant.

In Part IV we focus especially on four topics. The choice we made was a difficult one, because there have been many new developments in Markov chain theory over the last two decades. There is, therefore, a great deal of arbitrariness in these choices and important omissions. In Chapter 20, we assume that the state space is a complete separable metric space, but we no longer assume that the Markov chain is irreducible. Since it is no longer possible to construct an embedded regenerative process, the techniques of proof are completely different; the essential difference is that convergence in total variation distance may no longer hold, and it must be replaced by Wasserstein distances. We recall the main properties of these distances and in particular the duality theorem, which allows us to use coupling methods. We have essentially followed Hairer et al. (2011) in the geometric case and Butkovsky (2014) and Durmus et al. (2016) for the subgeometric case. However, the methods of proof and some of the results appear to be original. Chapter 21 covers central limit theorems of additive functions of Markov chains. The most direct approach is to use a martingale decomposition (with a remainder term) of the additive

functionals by introducing solutions of the Poisson equation. The approach is straightforward, and Poisson solutions exist under minimal technical assumptions (see Glynn and Meyn 1996), yet this method does not yield conditions close to optimal. A first approach to weaken these technical conditions was introduced in Kipnis and Varadhan (1985) and further developed by Maxwell and Woodroffe (2000): it keeps the martingale decomposition with remainder but replaces Poisson by resolvent solutions and uses tightness arguments. It yields conditions that are closer to being sufficient. A second approach, due to Gordin and Lifšic (1978) and later refined by many authors (see Rio 2017), uses another martingale decomposition and yields closely related (but nevertheless different) sets of conditions. We also discuss different expressions for the asymptotic variance, following Häggström and Rosenthal (2007). In Chapter 22, we discuss the spectral property of a Markov kernel P seen as an operator on an appropriately defined Banach space of complex functions and complex measures. We study the convergence to the stationary distribution using the particular structure of the spectrum of this operator; deep results can be obtained when the Markov kernel P is reversible (i.e., self-adjoint), as shown, for example, in Roberts and Tweedie (2001) and Kontoyiannis and Meyn (2012). We also introduce the notion of conductance and prove geometric convergence using conductance through Cheeger's inequalities, following Lawler and Sokal (1988) and Jarner and Yuen (2004). Finally, in Chapter 23 we give an introduction to sub-Gaussian concentration inequalities for Markov chains. We first show how McDiarmid's inequality can be extended to uniformly ergodic Markov kernels following Rio (2000a). We then discuss the equivalence between McDiarmid-type sub-Gaussian concentration inequalities and geometric ergodicity, using a result established in Dedecker and Gouëzel (2015). We finally obtain extensions of these inequalities for separately Lipschitz functions, following Djellout et al. (2004) and Joulin and Ollivier (2010).

We have chosen to illustrate the main results with simple examples. More substantial examples are considered in the exercises at the end of each chapter; the solutions of a majority of these exercises are provided. The reader is invited to use these exercises (which are mostly fairly direct applications of the material) to test their understanding of the theory. We have selected examples from different fields, including signal processing and automatic control, time-series analysis and Markov chain Monte Carlo simulation algorithms.

We do not cite bibliographical references in the body of the chapters, but we have added at the end of each chapter bibliographical indications. We give precise bibliographical indications for the most recent developments. For former results, we do not necessarily seek to attribute authorship to the original results. Meyn and Tweedie (1993b) covers in much greater detail the genesis of the earlier works.

The authors would like to thank the large number of people who at times contributed to this book. Alain Durmus, Gersende Fort, and François Roueff gave us valuable advice and helped us to clarify some of the derivations. Their contributions were essential. Christophe Andrieu, Gareth Roberts, Jeffrey Rosenthal and Alexander Veretennikov also deserve special thanks. They have been a very valuable source of inspiration for years.

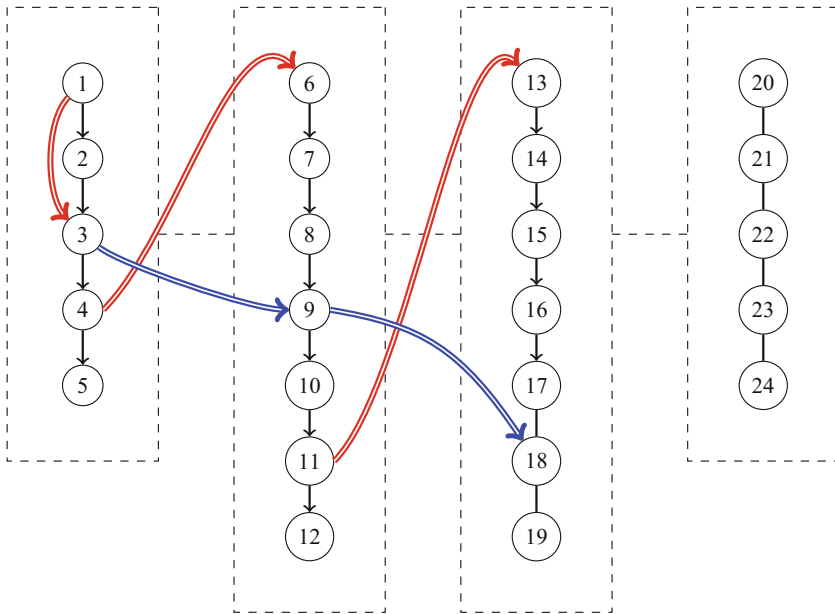


Fig. 1 Suggestion of playback order with respect to the different chapters of the book. The red arrows correspond to a possible path for a reader eager to focus only on the most fundamental results. The skipped chapters can then be investigated on a second reading. The blue arrows provide a fast track for a proof of the existence of an invariant measure and geometric rates of convergence for irreducible chains without the splitting technique. The chapters in the last part of the book are almost independent and can be read in any order.

We also benefited from the work of many colleagues who carefully reviewed parts of this book and helped us to correct errors and suggested improvements in the presentation: Yves Atchadé, David Barrera, Nicolas Brosse, Arnaud Doucet, Sylvain Le Corff, Matthieu Lerasle, Jimmy Olsson, Christian Robert, Claude Saint-Cricq, and Amandine Schreck.

We are also very grateful to all the students who for years helped us to polish what was at the beginning a set of rough lecture notes. Their questions and suggestions greatly helped us to improve the presentation and correct errors.

Évry, France
 Palaiseau, France
 Paris, France
 Nanterre, France

Randal Douc
 Eric Moulines
 Pierre Priouret
 Philippe Soulier

Contents

Part I Foundations

1	Markov Chains: Basic Definitions	3
1.1	Markov Chains	3
1.2	Kernels	6
1.3	Homogeneous Markov Chains	12
1.4	Invariant Measures and Stationarity	16
1.5	Reversibility	18
1.6	Markov Kernels on $L^P(\pi)$	20
1.7	Exercises	21
1.8	Bibliographical Notes	25
2	Examples of Markov Chains	27
2.1	Random Iterative Functions	27
2.2	Observation-Driven Models	35
2.3	Markov Chain Monte Carlo Algorithms	38
2.4	Exercises	49
2.5	Bibliographical Notes	51
3	Stopping Times and the Strong Markov Property	53
3.1	The Canonical Chain	54
3.2	Stopping Times	58
3.3	The Strong Markov Property	60
3.4	First-Entrance, Last-Exit Decomposition	64
3.5	Accessible and Attractive Sets	66
3.6	Return Times and Invariant Measures	67
3.7	Exercises	73
3.8	Bibliographical Notes	74

4	Martingales, Harmonic Functions and Poisson–Dirichlet Problems	75
4.1	Harmonic and Superharmonic Functions	75
4.2	The Potential Kernel	77
4.3	The Comparison Theorem	81
4.4	The Dirichlet and Poisson Problems	85
4.5	Time-Inhomogeneous Poisson–Dirichlet Problems	88
4.6	Exercises	89
4.7	Bibliographical Notes	95
5	Ergodic Theory for Markov Chains	97
5.1	Dynamical Systems	97
5.2	Markov Chain Ergodicity	104
5.3	Exercises	111
5.4	Bibliographical Notes	115
Part II Irreducible Chains: Basics		
6	Atomic Chains	119
6.1	Atoms	119
6.2	Recurrence and Transience	121
6.3	Period of an Atom	126
6.4	Subinvariant and Invariant Measures	128
6.5	Independence of the Excursions	134
6.6	Ratio Limit Theorems	135
6.7	The Central Limit Theorem	137
6.8	Exercises	140
6.9	Bibliographical Notes	144
7	Markov Chains on a Discrete State Space	145
7.1	Irreducibility, Recurrence, and Transience	145
7.2	Invariant Measures, Positive and Null Recurrence	146
7.3	Communication	148
7.4	Period	150
7.5	Drift Conditions for Recurrence and Transience	151
7.6	Convergence to the Invariant Probability	154
7.7	Exercises	159
7.8	Bibliographical Notes	164
8	Convergence of Atomic Markov Chains	165
8.1	Discrete-Time Renewal Theory	165
8.2	Renewal Theory and Atomic Markov Chains	175
8.3	Coupling Inequalities for Atomic Markov Chains	180
8.4	Exercises	187
8.5	Bibliographical Notes	189

- 9 Small Sets, Irreducibility, and Aperiodicity** 191
 - 9.1 Small Sets 191
 - 9.2 Irreducibility 194
 - 9.3 Periodicity and Aperiodicity 201
 - 9.4 Petite Sets 206
 - 9.5 Exercises 211
 - 9.6 Bibliographical Notes 215
 - 9.A Proof of Theorem 9.2.6 215
- 10 Transience, Recurrence, and Harris Recurrence** 221
 - 10.1 Recurrence and Transience 221
 - 10.2 Harris Recurrence 228
 - 10.3 Exercises 236
 - 10.4 Bibliographical Notes 239
- 11 Splitting Construction and Invariant Measures** 241
 - 11.1 The Splitting Construction 241
 - 11.2 Existence of Invariant Measures 247
 - 11.3 Convergence in Total Variation to the Stationary
Distribution 251
 - 11.4 Geometric Convergence in Total Variation Distance 253
 - 11.5 Exercises 258
 - 11.6 Bibliographical Notes 259
 - 11.A Another Proof of the Convergence of Harris Recurrent
Kernels 259
- 12 Feller and T -Kernels** 265
 - 12.1 Feller Kernels 265
 - 12.2 T -Kernels 270
 - 12.3 Existence of an Invariant Probability 274
 - 12.4 Topological Recurrence 277
 - 12.5 Exercises 279
 - 12.6 Bibliographical Notes 285
 - 12.A Linear Control Systems 285

Part III Irreducible Chains: Advanced Topics

- 13 Rates of Convergence for Atomic Markov Chains** 289
 - 13.1 Subgeometric Sequences 289
 - 13.2 Coupling Inequalities for Atomic Markov Chains 291
 - 13.3 Rates of Convergence in Total Variation Distance 303
 - 13.4 Rates of Convergence in f -Norm 305
 - 13.5 Exercises 311
 - 13.6 Bibliographical Notes 312

14	Geometric Recurrence and Regularity	313
14.1	f -Geometric Recurrence and Drift Conditions	313
14.2	f -Geometric Regularity	321
14.3	f -Geometric Regularity of the Skeletons	327
14.4	f -Geometric Regularity of the Split Kernel	332
14.5	Exercises	334
14.6	Bibliographical Notes	337
15	Geometric Rates of Convergence	339
15.1	Geometric Ergodicity	339
15.2	V -Uniform Geometric Ergodicity	349
15.3	Uniform Ergodicity	353
15.4	Exercises	356
15.5	Bibliographical Notes	358
16	(f, r)-Recurrence and Regularity	361
16.1	(f, r) -Recurrence and Drift Conditions	361
16.2	(f, r) -Regularity	370
16.3	(f, r) -Regularity of the Skeletons	377
16.4	(f, r) -Regularity of the Split Kernel	381
16.5	Exercises	382
16.6	Bibliographical Notes	383
17	Subgeometric Rates of Convergence	385
17.1	(f, r) -Ergodicity	385
17.2	Drift Conditions	392
17.3	Bibliographical Notes	399
17.A	Young Functions	399
18	Uniform and V-Geometric Ergodicity by Operator Methods	401
18.1	The Fixed-Point Theorem	401
18.2	Dobrushin Coefficient and Uniform Ergodicity	403
18.3	V -Dobrushin Coefficient	409
18.4	V -Uniformly Geometrically Ergodic Markov Kernel	412
18.5	Application of Uniform Ergodicity to the Existence of an Invariant Measure	415
18.6	Exercises	417
18.7	Bibliographical Notes	419
19	Coupling for Irreducible Kernels	421
19.1	Coupling	422
19.2	The Coupling Inequality	432
19.3	Distributional, Exact, and Maximal Coupling	435
19.4	A Coupling Proof of V -Geometric Ergodicity	441
19.5	A Coupling Proof of Subgeometric Ergodicity	444

19.6 Exercises 449
 19.7 Bibliographical Notes 451

Part IV Selected Topics

20 Convergence in the Wasserstein Distance 455
 20.1 The Wasserstein Distance 456
 20.2 Existence and Uniqueness of the Invariant Probability Measure 462
 20.3 Uniform Convergence in the Wasserstein Distance 465
 20.4 Nonuniform Geometric Convergence 471
 20.5 Subgeometric Rates of Convergence for the Wasserstein Distance 476
 20.6 Exercises 480
 20.7 Bibliographical Notes 485
 20.A Complements on the Wasserstein Distance 486

21 Central Limit Theorems 489
 21.1 Preliminaries 490
 21.2 The Poisson Equation 495
 21.3 The Resolvent Equation 503
 21.4 A Martingale Coboundary Decomposition 508
 21.5 Exercises 517
 21.6 Bibliographical Notes 519
 21.A A Covariance Inequality 520

22 Spectral Theory 523
 22.1 Spectrum 523
 22.2 Geometric and Exponential Convergence in $L^2(\pi)$ 530
 22.3 $L^p(\pi)$ -Exponential Convergence 538
 22.4 Cheeger’s Inequality 545
 22.5 Variance Bounds for Additive Functionals and the Central Limit Theorem for Reversible Markov Chains 553
 22.6 Exercises 560
 22.7 Bibliographical Notes 562
 22.A Operators on Banach and Hilbert Spaces 563
 22.B Spectral Measure 572

23 Concentration Inequalities 575
 23.1 Concentration Inequality for Independent Random Variables 576
 23.2 Concentration Inequality for Uniformly Ergodic Markov Chains 581
 23.3 Sub-Gaussian Concentration Inequalities for V -Geometrically Ergodic Markov Chains 587

- 23.4 Exponential Concentration Inequalities Under Wasserstein Contraction 594
- 23.5 Exercises 599
- 23.6 Bibliographical Notes 601
- Appendices** 603
- A Notations** 605
- B Topology, Measure and Probability** 609
 - B.1 Topology 609
 - B.2 Measures 612
 - B.3 Probability 618
- C Weak Convergence** 625
 - C.1 Convergence on Locally Compact Metric Spaces 625
 - C.2 Tightness 626
- D Total and V-Total Variation Distances** 629
 - D.1 Signed Measures 629
 - D.2 Total Variation Distance 631
 - D.3 V-Total Variation 635
- E Martingales** 637
 - E.1 Generalized Positive Supermartingales 637
 - E.2 Martingales 638
 - E.3 Martingale Convergence Theorems 639
 - E.4 Central Limit Theorems 641
- F Mixing Coefficients** 645
 - F.1 Definitions 645
 - F.2 Properties 646
 - F.3 Mixing Coefficients of Markov Chains 653
- G Solutions to Selected Exercises** 657
- References** 733
- Index** 753