

CrossMark



5 by David A. Levin and Yuval Peres

6 PROVIDENCE: AMERICAN MATHEMATICAL SOCIETY, 2017, XVI + 448 PP., 7 US \$84.00, ISBN 978-1-4704-2962-1

8 REVIEWED BY DAVID ALDOUS

1

10 nlike most books reviewed in the Mathematical 19 Intelligencer, this is definitely a textbook. It 12 assumes knowledge one might acquire in the first 13 two years of an undergraduate mathematics program-14 basic mathematical probability, plus linear algebra, a little 15 graph theory, and the infamous concept of "mathematical 16 maturity." It has the theorem-proof style of pure mathe-17 matics, but with friendly explanations of intuition and 18 motivation.

19 The Topic

20 The topic is one major aspect of what I like to call "modern 21 discrete probability," as opposed to balls-in-boxes-style 22 combinatorial probability. Here a Markov chain is best 23 envisaged as a particle jumping between a finite number of 24 states: from state *i*, it jumps to state *j* with some specified 25 probability p_{ii}—in other words, a random walk on an edge-26 weighted directed graph. A highlight of classical theory 27 says (omitting some details here and throughout) that 28 provided the graph is strongly connected, the time-t dis-29 tribution will converge as $t \to \infty$ to a limit distribution, the 30 stationary distribution. Theoretical mathematicians in the 31 1950s and 1960s focused on extending such general theory 32 to countably infinite or continuous state spaces, whereas 33 applied mathematicians at that time studied specific models 34 such as queueing and similar operations-research-style 35 models.

36 A different paradigm for finite-state chains emerged 37 around 1980. Consider a sequence of chains on state spaces 38 with increasing "size" parameter n, take some formaliza-39 tion of "close to stationarity," estimate the corresponding 40 "mixing time" τ_n (at which the time-*t* distribution is close to 41 stationarity), and study the behavior of τ_n as $n \to \infty$. In 42 other words, study size asymptotics, not time asymptotics. 43 A very simple example is to take binary strings of length n44 as states, and from a given state, we "jump" by switching a 45 random coordinate or (with small probability) not moving. 46 More popular examples arise from models of random card 47 shuffling: how many shuffles are needed to make an n-card 48 deck "well shuffled"?

49 But it turns out that there is a much wider range of 50 contexts in which mixing times are relevant. The notion of 51 a phase transition in statistical physics models such as the

	Dispatch :	31-8-2018	Pages : 2
5	□ LE ✓ CP		□ TYPESET ✓ DISK

52 Ising model corresponds to whether, in the natural asso-53 ciated Markov chain, τ_n increases polynomially or 54 exponentially in n. In using the Markov chain Monte Carlo (MCMC) method to reveal posterior distributions in com-55 plex big-data Bayesian statistics models, guaranteed 56 success depends on simulating for a number of steps larger 57 than the mixing time, and the mixing time cannot be 58 determined by simulation but requires theory. On the 59 mathematical side, there is an elegant and intuitively 60 helpful connection between electrical networks and 61 reversible (the matrix (p_{ij}) is symmetrizable) chains, 62 described previously in the 1984 undergraduate-level 63 64 monograph Random Walks and Electric Networks, by 65 Doyle and Snell.

As well as applications to these preexisting topics, the 66 new topic of approximate counting emerged in the 1980s. 67 As a basic example, one can readily use MCMC to 68 approximately sample uniformly from the set of colorings 69 of a given *n*-vertex graph (provided the number of avail-70 able colors is large enough), because one can bound the 71 relevant mixing time. What is less obvious at first sight 72 is that by then estimating probabilities recursively over 73 74 subgraphs of size n - 1, n - 2, ..., one can count approximately the number of colorings of the original graph. This 75 gives an example of a problem that is easily solved 76 approximately via a randomized algorithm but is difficult to 77 solve exactly by a deterministic algorithm. In the 2000s, this 78 line of enquiry led to a program to show that computa-79 tional complexity of optimization problems over random 80 data is related to phase transitions in corresponding statis-81 tical physics models, and this remains an active area of 82 mathematically deep research. 83

The Book

This is the expanded second edition (the first edition85appeared in 2009) of the first and only broad-ranging yet86carefully written textbook treatment of this topic. It has87been used in courses in many major universities, and has88ample exercises for students.89

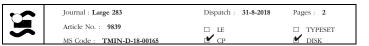
84

Part I of the book is "basics," illustrated by the follow-90 91 ing. One fundamental measure of "closeness to stationarity" uses variation distance, which involves the 92 worst-case (over events) additive error in approximating 93 94 the time-t distribution of a chain by the stationary probability. Upper-bounding the corresponding mixing time is 95 96 most easily done, where possible, via the coupling tech-97 nique. A conceptually different notion, mathematically natural in the reversible case, involves the relaxation time, 98 99 defined as 1 over the spectral gap of the symmetrized matrix (p_{ii}) . Techniques for upper-bounding this relaxation 100 time involve Dirichlet forms and the distinguished path 101 102 method. Also, properties of first hitting times and cover times (visiting all states) are non-obviously related to such 103 104 mixing times.

Part 2 continues to more advanced issues, including the motivating contexts alluded to before. Understanding mixing times in the Ising model in detail is still an active research topic, but Chapter 15 provides basic background. 108

© 2018 Springer Science+Business Media, LLC, part of Springer Nature

https://doi.org/10.1007/s00283-018-9839-x



110 Chains that are monotone with respect to a partial order 111 arise quite often and have special properties, described in 112 Chapter 22 (new). Various other special types of pro-113 cesses, such as lamplighter walks (Chapter 19) and the 114 exclusion process (Chapter 23, new), are treated carefully. 115 Also described are general theoretical developments by 116 Peres and coauthors over the last decade, such as the 117 "martingales and evolving sets" technique (Chapter 17) 118 and the intriguing connections between Cesàro mixing 119 times and hitting times on large subsets (Chapter 24, 120

new).
The twenty-six chapters comprising over four hundred
pages underline the vast range of contexts and techniques

123 being discussed.

The Bottom Line

Taking a recently emerged topic with a massive research literature and writing a textbook that can take a student from125basic undergraduate mathematics to the ability to read current126research papers is a hugely impressive achievement. This128book will long remain the definitive required reading for129anyone wishing to engage the topic more than superficially.130

124

		131
		132
Department of Statistics		133
367 Evans Hall		134
Berkeley, CA 94720	·	135
USA		136
e-mail: aldous@stat.berkelev.edu		137

THE MATHEMATICAL INTELLIGENCER