Continuous Time Markov Processes
An Introduction

Thomas M. Liggett

Graduate Studies in Mathematics
Volume 113

American Mathematical Society

Continuous Time Markov Processes

An Introduction

Thomas M. Liggett

Graduate Studies
in Mathematics
Volume 113

American Mathematical Society
Providence, Rhode Island

This work was supported in part by NSF grant DMS-0301795

2000 *Mathematics Subject Classification.* Primary 60J25, 60J27, 60J65; Secondary 35J05, 60J35, 60K35.

For additional information and updates on this book, visit
www.ams.org/bookpages/gsm-113

Library of Congress Cataloging-in-Publication Data

Liggett, Thomas M. (Thomas Milton), 1944–
Continuous time Markov processes : an introduction / Thomas M. Liggett.
p. cm. — (Graduate studies in mathematics ; v. 113)
Includes bibliographical references and index.
ISBN 978-0-8218-4949-1 (alk. paper)
1. Markov processes. 2. Stochastic integrals. I. Title.

QA274.7.L54 2010
519.2′33—dc22

2009045839

---

10 9 8 7 6 5 4 3 2 1 15 14 13 12 11 10

Continuous Time Markov Processes

An Introduction

Preface

Students are often surprised when they first hear the following definition: A stochastic process is a collection of random variables indexed by time. There seems to be no content here. There is no structure. How can anyone say anything of value about a stochastic process? The content and structure are in fact provided by the definitions of the various classes of stochastic processes that are so important for both theory and applications. There are processes in discrete or continuous time. There are processes on countable or general state spaces. There are Markov processes, random walks, Gaussian processes, diffusion processes, martingales, stable processes, infinitely divisible processes, stationary processes, and many more. There are entire books written about each of these types of stochastic process.

The purpose of this book is to provide an introduction to a particularly important class of stochastic processes — continuous time Markov processes. My intention is that it be used as a text for the second half of a year-long course on measure-theoretic probability theory. The first half of such a course typically deals with the classical limit theorems for sums of independent random variables (laws of large numbers, central limit theorems, random infinite series), and with some of the basic discrete time stochastic processes (martingales, random walks, stationary sequences). Alternatively, the book can be used in a semester-long special topics course for students who have completed the basic year-long course. In this case, students will probably already be familiar with the material in Chapter 1, so the course would start with Chapter 2.

The present book stresses the new issues that appear in continuous time. A difference that arises immediately is in the definition of the process. A discrete time Markov process is defined by specifying the law that leads from

the state at one time to that at the next time. This approach is not possible in continuous time. In most cases, it is necessary to describe the transition law infinitesimally in time, and then prove under appropriate conditions that this description leads to a well-defined process for all time.

We begin with an introduction to Brownian motion, which is certainly the most important continuous time stochastic process. It is a special case of many of the types listed above — it is Markov, Gaussian, a diffusion, a martingale, stable, and infinitely divisible. It plays a fundamental role in stochastic calculus, and hence in financial mathematics. Through Donsker's theorem, it provides a framework for far reaching generalizations of the classical central limit theorem. While we will concentrate on this one process in Chapter 1, we will also discuss there the extent to which results and techniques apply (or do not apply) more generally. The infinitesimal definition mentioned in the previous paragraph is not necessary in the case of Brownian motion. However, our discussion of Brownian motion sets the stage for the setup that is required for processes that are defined in that way.

Next we discuss the construction problem for continuous time Markov chains. (The word "chain" here refers to the countability of the state space.) The main issue is to determine when the infinitesimal description of the process (given by the Q-matrix) uniquely determines the process via Kolmogorov's backward equations.

With an understanding of these two examples — Brownian motion and continuous time Markov chains — we will be in a position to consider the issue of defining the process in greater generality. Key here is the Hille-Yosida theorem, which links the infinitesimal description of the process (the generator) to the evolution of the process over time (the semigroup). Since usually only the generator is known explicitly, we will discuss how one deduces properties of the process from information about the generator. The main examples at this point are variants of Brownian motion, in which the relative speed of the particle varies spatially, and/or there is a special behavior at the boundary of the state space.

As an application of the theory of semigroups and generators, we then provide an introduction to a somewhat more recently developed area of probability theory — interacting particle systems. This is a class of probabilistic models that come up in many areas of application — physics, biology, computer science, and even a bit in economics and sociology. Infinitely many agents evolve in time according to certain probabilistic rules that involve interactions among the agents. The nature of these rules is dictated by the area of application. The main issue here is the nature of the long time behavior of the process.

Next we give an introduction to stochastic integration with respect to Brownian motion and other continuous (semi)martingales. Not only is this an important probabilistic tool, but in recent years, it has become an essential part of financial mathematics. We define the Itô integral and study its properties, which are quite different from those of ordinary integrals, as a consequence of the lack of smoothness of Brownian paths. Then we use it to construct local time for Brownian motion, and apply it to give a new perspective on some of the Brownian relatives from Chapter 3.

In the final chapter, we return to Brownian motion, now in higher dimensions, and describe one of its great successes in analysis — that of providing a probabilistic solution to the classical Dirichlet problem. This problem asks for harmonic functions (those satisfying $\Delta h = 0$) in a domain in R^n with prescribed boundary values. Then we discuss the Poisson equation $\frac{1}{2}\Delta h = -f$. Solutions to the Dirichlet problem and Poisson equation provide concrete answers to many problems involving Brownian motion in R^n. Examples are exit distributions from domains, and expected occupation times of subsets prior to exiting a domain.

The prerequisite for reading this book is a semester course in measure-theoretic probability that includes the material in the first four chapters of [**18**], for example. In particular, students should be familiar with laws of large numbers, central limit theorems, random walks, the basics of discrete time Markov chains, and discrete time martingales. To facilitate referring to this material, I have included the main definitions and results (mostly without proofs) in the Appendix. Approximately 200 exercises are placed within the sections as the relevant material is covered.

Chapters 1 and 2 are largely independent of one another, but should be read before Chapter 3. They provide motivation for the more abstract treatment of Feller processes there. The main places where Chapter 2 relies on material from Chapter 1 are in the discussions of the Markov and strong Markov properties. Rather than prove these in some generality, our approach is to prove them in the concrete context of Brownian motion. By making explicit the properties of Brownian motion that are used in the proofs, we are able simply to refer back to those proofs when these properties are discussed in Chapters 2 and 3.

The hearts of Chapters 2 and 3 are Sections 2.5 and 3.3 respectively. The prior sections in these chapters are intended to provide motivation for the transition from infinitesimal description to time evolution that is explained in those sectons. Therefore, the earlier sections need not be covered in full detail. In my classes, I often state the main results from the earlier sections without proving many of them, in order to allow ample time for the more important material in Sections 2.5 and 3.3.

The last three chapters can be covered in any order. Chapters 5 and 6 rely only slightly on Chapters 2 and 3, so one can easily create a short course based on Chapters 1, 5 and 6.

This book is based on courses I have taught at UCLA over many years. Unlike many universities, UCLA operates on the quarter system. I have typically covered most of the material in Chapters 1–3 and 6 in the third quarter of the graduate probability course, and Chapters 4 and 5 in special topics courses. There is more than enough material here for a semester course, even if Chapter 1 is skipped because students are already familiar with one-dimensional Brownian motion.

Despite my best efforts, some errors have probably made their way into the text. I will maintain a list of corrections at

http://www.math.ucla.edu/~tml/

Readers are encouraged to send me corrections at tml@math.ucla.edu.

As is usually the case with a text of this type, I have benefitted greatly from the work of previous authors, including those of [**12**], [**18**], [**21**], [**22**], [**39**], and [**40**]. I appreciate the comments and corrections provided by P. Caputo, S. Roch, and A. Vandenberg-Rodes, and especially T. Richthammer and F. Zhang, who read much of this book very carefully.

Thomas M. Liggett

Contents

Chapter 1

One-Dimensional Brownian Motion

1.1. Some motivation

The biologist Robert Brown noticed almost two hundred years ago that bits of pollen suspended in water undergo chaotic behavior. The bits of pollen are much more massive than the molecules of water, but of course there are many more of these molecules than there are bits of pollen. The chaotic motion of the pollen is the result of many infinitesimal jolts by the water molecules. By the central limit theorem (CLT), the law of the motion of the pollen should be closely related to the normal distribution. We now call this law Brownian motion.

During the past half century or so, Brownian motion has turned out to be a very versatile tool for both theory and applications. As we will see in Chapter 6, it provides a very elegant and general treatment of the Dirichlet problem, which asks for harmonic functions on a domain with prescribed boundary values. It is also the main building block for the theory of stochastic calculus, which is the subject of Chapter 5. Via stochastic calculus, it has played an important role in the development of financial mathematics.

As we will see later in this chapter, Brownian paths are quite rough — they are of unbounded variation in every time interval. Therefore, integrals with respect to them cannot be defined in the Stieltjes sense. A new type of integral must be defined, which carries the name of K. Itô, and more recently, of W. Doeblin. This new integral has some unexpected properties. Here is an example: If $B(t)$ is standard Brownian motion at time t with

$B(0) = 0$, then

$$\int_0^t B(s)\, dB(s) = \frac{1}{2}\big[B^2(t) - t\big]. \tag{1.1}$$

Of course, if $B(t)$ could be used as an integrator in the Stieltjes sense, and this were the Stieltjes integral, the right side would not contain the term $-t$.

There are also many important applications of Brownian motion connected with the classical limit theorems of probability theory. If $\xi_1, \xi_2, \ldots$ are i.i.d. random variables with mean zero and variance one, and

$$S_n = \xi_1 + \cdots + \xi_n$$

are their partial sums, the CLT says that $S_n/\sqrt{n}$ converges in distribution to the standard normal law. How can one embed the CLT into a more general theory that includes as one of its consequences the fact that $\max\{0, S_1, \ldots, S_n\}/\sqrt{n}$ converges in distribution to the absolute value of a standard normal? The answer involves Brownian motion in a crucial way, as we will see later in this chapter. Here is an early hint: For $t \geq 0$ and $n \geq 1$, let

$$X_n(t) = \frac{S_{[nt]}}{\sqrt{n}}, \tag{1.2}$$

where $[\cdot]$ is the integer part function. Then $X_n(1) = S_n/\sqrt{n}$, and

$$\max_{0 \leq t \leq 1} X_n(t) = \frac{\max\{0, S_1, \ldots, S_n\}}{\sqrt{n}}.$$

So, we have written both functionals of the partial sums in terms of the stochastic process $X_n(t)$. Once we show that X_n converges in an appropriate sense to Brownian motion, we will have a limit theorem for

$$\max\{0, S_1, \ldots, S_n\},$$

as well as for many other functions of the partial sums.

This chapter represents but a very small introduction to a huge field. For further reading, see [**35**] and [**40**].

1.2. The multivariate Gaussian distribution

Before defining Brownian motion, we will need to review the multivariate Gaussian distribution. Recall that a random variable ξ has the standard Gaussian (or normal) distribution $N(0,1)$ if it has density

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty.$$

It is said to be univariate Gaussian if it can be written in the form $\xi = a\zeta + b$, where ζ is standard Gaussian and a, b are real. Note that this definition

allows ξ to have zero variance. The Gaussian distribution with mean m and variance σ^2 (obtained above if $b = m$ and $a^2 = \sigma^2$) is denoted by $N(m, \sigma^2)$.

Definition 1.1. The real random vector $(\xi_1, \ldots, \xi_n)$ is said to be multivariate Gaussian if all linear combinations

$$\sum_{k=1}^{n} a_k \xi_k$$

of the components have univariate Gaussian distributions.

Remark 1.2. (a) If $\xi_1, \ldots, \xi_n$ are independent Gaussians, then $(\xi_1, \ldots, \xi_n)$ is multivariate Gaussian.

(b) Definition 1.1 is much stronger than the statement that each ξ_k is Gaussian. For example, suppose ζ is standard Gaussian, and

$$\xi = \begin{cases} +\zeta & \text{if } |\zeta| \le 1; \\ -\zeta & \text{if } |\zeta| > 1. \end{cases}$$

Then ξ is also standard Gaussian. However, since $|\zeta + \xi| \le 2$ and $\zeta + \xi$ is not constant, $\zeta + \xi$ is not Gaussian, so (ζ, ξ) is not bivariate Gaussian.

Remark 1.3. Definition 1.1 has a number of advantages over the alternative, in which one specifies the joint density of $(\xi_1, \ldots, \xi_n)$:

(a) It does not require that $(\xi_1, \ldots, \xi_n)$ have a density. For example, (ξ, ξ) is bivariate Gaussian if ξ is Gaussian.

(b) It makes the next result immediate.

Proposition 1.4. *Suppose $\xi = (\xi_1, \ldots, \xi_n)$ is Gaussian and A is an $m \times n$ matrix. Then the random vector $\zeta = A\xi$ is also Gaussian.*

Proof. Any linear combination of $\zeta_1, \ldots, \zeta_m$ is some other linear combination of $\xi_1, \ldots, \xi_n$. □

An important property of a multivariate Gaussian vector ξ is that its distribution is determined by the mean vector $E\xi$ and the covariance matrix, whose (i, j) entry is $\mathrm{Cov}(\xi_i, \xi_j)$. To check this statement, we use characteristic functions. Recall that the characteristic function of a random variable with the $N(m, \sigma^2)$ distribution is

$$\exp\left\{itm - \frac{1}{2}t^2\sigma^2\right\}.$$

Therefore, if $\xi = (\xi_1, \ldots, \xi_n)$ is multivariate Gaussian, its joint characteristic function is given by

$$\phi(t_1, \ldots, t_n) = E \exp\left\{i \sum_{j=1}^{n} t_j \xi_j\right\} = \exp\left\{im - \frac{1}{2}\sigma^2\right\},$$

where m and σ^2 are the mean and variance of $\sum_{j=1}^n t_j \xi_j$:

$$m = \sum_{j=1}^n t_j E\xi_j \quad \text{and} \quad \sigma^2 = \sum_{j,k=1}^n t_j t_k \operatorname{Cov}(\xi_j, \xi_k).$$

Since $\phi(t_1, \ldots, t_n)$ depends on ξ only through its mean vector and covariance matrix, these determine the characteristic function of ξ, and hence its distribution by Proposition A.24. This observation has the following consequence:

Proposition 1.5. *If $\xi = (\xi_1, \ldots, \xi_n)$ is multivariate Gaussian, then the random variables $\xi_1, \ldots, \xi_n$ are independent if and only if they are uncorrelated.*

Proof. That independence implies uncorrelatedness is always true for random variables with finite second moments. For the converse, suppose that $\xi_1, \ldots, \xi_n$ are uncorrelated, i.e., that $\operatorname{Cov}(\xi_j, \xi_k) = 0$ for $j \neq k$. Take $\zeta_1, \ldots, \zeta_n$ to be independent, with ζ_i having the same distribution as ξ_i. Then ξ and $\zeta = (\zeta_1, \ldots, \zeta_n)$ have the same characteristic function, and hence the same distribution, by Proposition A.24. It follows that $\xi_1, \ldots, \xi_n$ are independent. □

The next exercise will be useful in Chapter 6 — see Proposition 6.12.

Exercise 1.6. Show that if $\xi = (\xi_1, \ldots, \xi_n)$, where $\xi_1, \ldots, \xi_n$ are i.i.d. standard Gaussian random variables, and O is an $n \times n$ orthogonal matrix, then $O\xi$ has the same distribution as ξ.

For the next exercise, recall that $\Rightarrow$ denotes convergence in distribution — see Definition A.18.

Exercise 1.7. (a) Suppose that $\xi_k \Rightarrow \xi$ and that ξ_k has the $N(m_k, \sigma_k^2)$ distribution for each k. Prove that ξ is $N(m, \sigma^2)$ for some m and σ^2, and that $m_k \to m$ and $\sigma_k^2 \to \sigma^2$. (Suggestion: First reduce the problem to the mean zero case by symmetrization, i.e., consider $\xi_k - \overline{\xi}_k$, where $\overline{\xi}_k$ is independent of ξ_k and has the same distribution.)

(b) State an analogue of (a) for Gaussian random vectors, and prove it using part (a). (Recall the Cramér-Wold device, Theorem A.26.)

The main topic of this book is a class of stochastic processes; in this chapter, they are Gaussian. We conclude this section with formal definitions of these concepts.

Definition 1.8. A stochastic process is a collection of random variables indexed by time. It is a discrete time process if the index set is a subset of the integers, and a continuous time process if the index set is $[0, \infty)$ (or sometimes, $(-\infty, \infty)$).

Definition 1.9. A stochastic process $X(t)$ is Gaussian if for any $n \geq 1$ and any choice of times $t_1, \ldots, t_n$, the random vector $(X(t_1), \ldots, X(t_n))$ has a multivariate Gaussian distribution. Its mean and covariance functions are $EX(t)$ and $\mathrm{Cov}(X(s), X(t))$ respectively.

1.3. Processes with stationary independent increments

As we will see shortly, Brownian motion is not only a Gaussian process, but is a process with two other important properties — stationarity and independence of its increments. Here is the relevant definition.

Definition 1.10. A stochastic process $(X(t), t \geq 0)$ has stationary increments if the distribution of $X(t) - X(s)$ depends only on $t - s$ for any $0 \leq s \leq t$. It has independent increments if the random variables $\{X(t_{j+1}) - X(t_j), 1 \leq j < n\}$ are independent whenever $0 \leq t_1 < t_2 < \cdots < t_n$ and $n \geq 1$.

The simplest process with stationary independent increments is the Poisson process $N(t)$ with parameter $\lambda > 0$. It has the following properties:

(i) $N(t, \omega)$ is an increasing right continuous step function in t with jumps of size 1,

and

(ii) $N(t) - N(s)$ is Poisson distributed with parameter $\lambda(t - s)$ for $0 \leq s < t$.

It can be constructed in the following way: Let $\tau_1, \tau_2, \ldots$ be independent and identically distributed random variables that are exponentially distributed with parameter λ. Then let

$$N(t) = \#\{k \geq 1 : \tau_1 + \cdots + \tau_k \leq t\}. \tag{1.3}$$

Exercise 1.11. With $N(t)$ defined as in (1.3), show that if $0 < s < t$, then $N(s)$ and $N(t) - N(s)$ are independent Poisson distributed random variables with parameters λs and $\lambda(t - s)$ respectively.

1.4. Definition of Brownian motion

To see that the properties introduced in the previous two sections are likely to have a bearing on the definition of Brownian motion, note that the process $X_n(t)$ defined in (1.2) has independent increments, and that except for the effect of time discretization, it has stationary increments. Therefore, any limit $X(t)$ of $X_n(t)$ as $n \to \infty$, if it exists in any reasonable sense, will have stationary independent increments. Also, by the central limit theorem, $X(t)$ will have the $N(0, t)$ distribution. Thus, we would expect Brownian motion

to be Gaussian and have stationary independent increments. The following result relates these properties.

Proposition 1.12. *The following two statements are equivalent for a stochastic process* $(X(t), t \geq 0)$*:*

(a) $X(t)$ *has stationary independent increments, and* $X(t)$ *is* $N(0,t)$ *for each* $t \geq 0$*.*

(b) $X(t)$ *is a Gaussian process with* $EX(t) = 0$ *and*

$$\mathrm{Cov}(X(s), X(t)) = s \wedge t.$$

Proof. Suppose (a) holds. To show that the process is Gaussian, take a_k's and t_k's as required in Definitions 1.1 and 1.9. Without loss of generality, we may assume that $0 = t_0 < t_1 < \cdots < t_n$. Summing by parts, and using $X(0) = 0$, we see that there are b_k's so that

$$\sum_{k=1}^{n} a_k X(t_k) = \sum_{k=1}^{n} b_k \big[X(t_k) - X(t_{k-1})\big].$$

The right side is a sum of independent Gaussians, and hence is Gaussian. To check the covariance statement, take $s < t$ and write

$$\begin{aligned} \mathrm{Cov}(X(s), X(t)) &= EX(s)X(t) \\ &= EX(s)\big[X(t) - X(s)\big] + EX^2(s) \\ &= s = s \wedge t. \end{aligned}$$

For the converse, assume (b). Then for $s < t$, $X(t) - X(s)$ is Gaussian with mean zero and

$$\mathrm{Var}(X(t) - X(s)) = t - 2(s \wedge t) + s = t - s,$$

so the process has stationary increments and has the correct marginal distributions.

To check independence of the increments, take $0 \leq t_1 < t_2 < \cdots < t_n$, and write the vector of increments in the form

$$(X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})) = A(X(t_1), \ldots, X(t_n))$$

for an appropriately chosen matrix A. Therefore, by Proposition 1.4, the vector of increments is Gaussian. So, in order to check the independence of the increments, it is enough by Proposition 1.5 to show that the increments are uncorrelated. To do so, take $u < v \leq s < t$. Then

$$\begin{aligned} \mathrm{Cov}(X(v) - X(u), X(t) - X(s)) &= v \wedge t - v \wedge s - u \wedge t + u \wedge s \\ &= v - v - u + u = 0. \end{aligned}$$

□