Continuous Time Markov Processes: An Introduction

Thomas M. Liggett

Department of Mathematics, UCLA

Current address: Department of Mathematics, University of California, Los Angeles CA 90095

E-mail address: tml@math.ucla.edu

1991 *Mathematics Subject Classification.* Primary 60J25, 60J27, 60J65;
Secondary 35J05, 60J35, 60K35

Key words and phrases. Probability theory, Brownian motion, Markov chains, Feller processes, the voter model, the contact process, exclusion processes, stochastic calculus, Dirichlet problem

This work was supported in part by NSF Grant #DMS-0301795.

Abstract. This is a textbook intended for use in the second semester of the basic graduate course in probability theory and/or in a semester topics course to follow the one year course.

Contents

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. 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 discussed in the previous paragraph is not necessary in the case of Brownian motion; however it sets the stage for the set up 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 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 calculus 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 the Brownian relatives from Chapter 3.

In the final chapter, we return to Brownian motion, and describe one of its great successes in analysis – that of providing a solution to the classical Dirichlet problem, which 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. Over 150 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. 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 will be 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. Therefore, the earlier sections need not be covered in full detail. I often state the main results from the earlier sections without proving many of them, in order to allow more time for the transition from infinitesimal description to time evolution. The last three chapters can be covered in any order.

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.

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**], [**20**], [**21**], [**22**], [**27**], [**35**], [**37**], [**38**], and [**42**]. I also appreciate the comments and corrections provided by P. Caputo, S. Roch, and A. Vandenberg-Rodes, and especially T. Richthammer, who read much of the book very carefully.

Thomas M. Liggett

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, 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 Brownian motion at time t, 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 have the term $-t$ in it.

There are also the important applications connected with the classical limit theorems of probability theory. If $\xi_1, \xi_2, ...$ are i.i.d. random variables with mean zero and variance one and $S_n = \xi_1 + \cdots + \xi_n$, the CLT says that $S_n/\sqrt{n}$ converges in distribution to the standard normal. How can one embed the CLT into a more general theory that includes as one of its consequences the fact that $\max\{0, S_1, ..., S_n\}/\sqrt{n}$ 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

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

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

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

as well as for many other functions of the S_n's.

This chapter represents but a very small introduction to a huge field. For further reading, see [**34**] and [**38**].

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 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 $(\xi_1, ..., \xi_n)$ is said to be multivariate Gaussian if all linear combinations

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

have univariate Gaussian distributions.

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

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

(a) It does not require that $(\xi_1, ..., \xi_n)$ 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 $\xi = (\xi_1, ..., \xi_n)$ 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 $\zeta_1, ..., \zeta_m$ is some other linear combination of $\xi_1, ..., \xi_n$. □

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 $Cov(\xi_i, \xi_j)$. 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 $\xi = (\xi_1, ..., \xi_n)$ is multivariate Gaussian, its joint characteristic function is given by

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

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

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

Since $\phi(t_1, ..., t_n)$ 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.22. This observation has the following consequences:

Proposition 1.5. *If $\xi = (\xi_1, ..., \xi_n)$ is multivariate Gaussian, then $\xi_1, ..., \xi_n$ 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 $\xi_1, ..., \xi_n$ are uncorrelated, i.e., that $Cov(\xi_j, \xi_k) = 0$ for $j \neq k$. Take $\zeta_1, ..., \zeta_n$ to be independent, with ζ_i having the same distribution as ξ_i. Then ξ and $\zeta = (\zeta_1, ..., \zeta_n)$ have the same characteristic function, and hence the same distribution, by Proposition A.22. It follows that $\xi_1, ..., \xi_n$ are independent. □

Exercise 1.6. Show that if $\xi = (\xi_1, ..., \xi_n)$, where $\xi_1, ..., \xi_n$ 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.

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 $\sigma_k^2 \to \sigma^2$.

(b) State and prove an analogue of (a) for Gaussian random vectors.

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 $\{0, 1, 2, ...\}$, and a continuous time process if the index set is $[0, \infty)$.

Definition 1.9. A stochastic process $X(t)$ is Gaussian if for any $n \geq 1$ and any choice of times $t_1, ..., t_n$, the random vector $(X(t_1), ..., X(t_n))$ has a multivariate Gaussian distribution. Its mean and covariance functions are $EX(t)$ and $Cov(X(s), X(t))$ 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 properties that $N(t, \omega)$ is an increasing right continuous step function in t with jumps of size 1, and $N(t) - N(s)$ is Poisson distributed with parameter $\lambda(t - s)$ for $0 \leq s < t$. It can be constructed in the following way: Let $\tau_1, \tau_2, ...$ be independent and identically distributed random variables that are exponentially distributed with parameter λ. Then let

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

1.4. Definition of Brownian motion

To see that the properties introduced in the previous two sections may have a bearing on our definition of Brownian motion, note that the process $X_n(t)$ defined in (1.2) has independent increments, and except for the effect of time discretization, 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 1.11. *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*

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

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 is hence Gaussian. To check the covariance statement, take $s < t$ and write

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

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

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

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

If $0 \le t_1 < t_2 < \cdots < t_n$, the vector of increments can be written in the form

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

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 $u < v \le s < t$. Then

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

□

Exercise 1.12. Suppose $X(t)$ is a stochastic process with stationary independent increments that satisfies $EX(1) = 0, EX^2(1) = 1$, and $X(t)$ has the same distribution as $\sqrt{t}X(1)$ for $t \ge 0$. Show that $X(t) - X(s)$ is $N(0, |t-s|)$ for $s, t \ge 0$.

Definition 1.13. A stochastic process $(X(t), t \ge 0)$, is said to have continuous paths if

$$P(\{\omega : X(t, \omega) \text{ is continuous in } t\}) = 1. \tag{1.3}$$

Definition 1.14. Standard Brownian motion $B(t)$ is a stochastic process with continuous paths that satisfies the equivalent properties (a) and (b) in Proposition 1.11.

Of course, it is not at all obvious that there exists a probability space on which one can construct a standard Brownian motion. Showing that this is the case is the objective of the next section. Taking this for granted for the time being, here are two exercises that provide some practice with the definition. In both cases, $B(t)$ is standard Brownian motion.

Exercise 1.15. Let

$$X(t) = \int_0^t B(s)ds.$$

(a) Explain why $X(t)$ is a Gaussian process.

(b) Compute the mean and covariance functions of X.

(c) Compute $E(X(t) - X(s))^2$, and compare its rate of decay as $t \downarrow s$ with that of $E(B(t) - B(s))^2$.

Exercise 1.16. Compute

$$P(B(s) > 0, B(t) > 0), \quad 0 < s < t.$$

(Recall Exercise 1.6.)

If it were not for the path continuity requirement in Definition 1.14, the existence of standard Brownian motion on some probability space would follow from Kolmogorov's extension theorem – see Theorem A.1. Since every event in this probability space is determined by the process at only countably many times, and continuity of a path is not determined by its values at countably many times, this approach would lead to the awkward situation in which the set in question, $C = \{\omega : B(t, \omega) \text{ is continuous in } t\}$, is not an event. Therefore, we would not even be able to discuss the issue of whether its probability is 1.

The situation is even more serious than this. Even if C were measurable, it would not be possible to prove that $P(C) = 1$ follows from properties (a) and (b). To see this, take a process B on some probability space satisfying properties (a) and (b), and let τ be a continuous random variable that is independent of B. Define a new process by

$$X(t, \omega) = \begin{cases} B(t, \omega) & \text{if } t \neq \tau(\omega); \\ B(t, \omega) + 1 & \text{if } t = \tau(\omega). \end{cases}$$

Then not both B and X can have continuous paths. However, since

$$P(X(t) = B(t)) = P(\tau \neq t) = 1$$

for every t, it follows that X also satisfies properties (a) and (b) in Proposition 1.11. Therefore, there exist processes that satisfy properties (a) and (b) but do not have continuous paths. Note that

$$P(X(t) = B(t) \text{ for all } t) = 0.$$

The fact that this can happen even though $P(X(t) = B(t)) = 1$ for every t is an early indication of how different things can be in discrete and continuous time. In the next section, we will see how to get around these difficulties.

These comments suggest the importance of the following definition.

Definition 1.17. Two stochastic processes $X(t)$ and $Y(t)$ are versions of one another it $P(X(t) = Y(t)) = 1$ for all t.

The following exercise introduces an important variant of Brownian motion, which is known as the Brownian bridge, or tied down Brownian motion. It arises in the study of empirical distribution functions.

Exercise 1.18. Define $X(t) = B(t) - tB(1)$ for $0 \le t \le 1$, where B is standard Brownian motion.

(a) Show that X is a Gaussian process, and compute its covariance function $Cov(X(s), X(t))$.

(b) Show that for $0 < t_1 < \cdots < t_n < 1$, the (joint) distribution of $\big((B(t_1), ..., B(t_n)) \mid |B(1)| \le \epsilon\big)$ converges to the (joint) distribution of $(X(t_1), ..., X(t_n))$ as $\epsilon \downarrow 0$.

The next exercise gives a limit law for the occupation time of A by Brownian motion up to time t – take the f below to be 1_A. This approach is computationally intensive, but provides good practice in working with Gaussian integrals and basic properties of Brownian motion. The proof becomes much easier and neater once we have developed some theory – see Exercise 5.54.

Exercise 1.19. Let B be standard Brownian motion, and put

$$X(t) = \frac{1}{\sqrt{t}} \int_0^t f(B(s))ds,$$

where $f \in L_1(R^1)$ and $\int f(x)dx = 1$.

(a) Show that

$$\lim_{t\to\infty} EX(t) = \sqrt{2/\pi} \quad \text{and} \quad \lim_{t\to\infty} EX^2(t) = 1.$$

(b) Use the method of moments – see Theorem A.27 – to prove that $X(t) \Rightarrow |Z|$, where Z is standard normal, following the outline below.

(i) For $\alpha_1, ..., \alpha_k > 0$, let $I(\alpha_1, ..., \alpha_k) =$

$$\int \cdots \int_{0<r_1<\cdots<r_k<1} r_1^{\alpha_1 - 1}(r_2 - r_1)^{\alpha_2 - 1} \cdots (r_k - r_{k-1})^{\alpha_k - 1} dr_1 \cdots dr_k.$$

Show that

$$\lim_{t\to\infty} EX^k(t) = \frac{k!}{(2\pi)^{k/2}} I\left(\frac{1}{2}, ..., \frac{1}{2}\right).$$

(ii) Using the Beta integral

$$\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},$$

check that

$$I(\alpha_1, ..., \alpha_k) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} I(\alpha_1+\alpha_2, \alpha_3, ..., \alpha_k),$$

so that

$$I\left(\frac{1}{2}, ..., \frac{1}{2}\right) = \frac{[\Gamma(1/2)]^k}{\Gamma(k/2)(k/2)}.$$

(iii) Check the recursion

$$E|Z|^k = (k-1)E|Z|^{k-2}, \quad k \geq 2,$$

and use it to compute $E|Z|^k$.

1.5. The construction

In this section, we will give one construction of Brownian motion. Another construction is outlined in the exercises.

Theorem 1.20. *There exists a probability space* $(\Omega, \mathcal{F}, P)$ *on which standard Brownian motion* B *exists.*

Proof. Properties (a) and (b) in Proposition 1.11 specify the finite dimensional distributions of B. By Kolmogorov's extension theorem, Theorem A.1, there exists a probability space on which the random variables $B(t)$ are defined for $t \in Q^+$, the set of positive rationals. We will prove that for every $N \geq 1$,

(1.4) $\quad B(t, \omega)$ is uniformly continuous in t for $t \in Q \cap [0, N]$ *a.s.*

Once this is done, $B(t)$ can be extended to all $t \geq 0$ by continuity. Note that the uniformity is important here. If $b \notin Q$, the function

$$f(t) = \begin{cases} 1 & \text{if } t \geq b; \\ 0 & \text{if } t < b \end{cases}$$

is continuous on Q, but not uniformly continuous on Q. It cannot be extended from Q to R^1 by continuity.

Let

$$\Delta_n = \sup_{\substack{s,t \in Q \cap [0,N] \\ |s-t| \leq \frac{1}{n}}} |B(t) - B(s)|.$$

We need to prove that $\Delta_n \to 0$ a.s. Since Δ_n is decreasing in n, it is enough to prove convergence in probability. To see this, recall that convergence in probability implies that a subsequence converges a.s. By the monotonicity, convergence along a subsequence implies convergence along the full sequence.

The next step is to reduce the number of arguments of B that need to be considered in the supremum. To do so, let

$$Y_{k,n} = \sup_{\substack{t\in Q \\ \frac{k-1}{n}\leq t\leq \frac{k}{n}}} \left| B(t) - B\left(\frac{k-1}{n}\right)\right|.$$

Then

$$\Delta_n \leq 3 \max_{1\leq k\leq nN} Y_{k,n}. \tag{1.5}$$

The factor of 3 above arises in the following way. For a given t, choose k so that $\frac{k-1}{n} \leq t \leq \frac{k}{n}$. If $|s-t| \leq \frac{1}{n}$, then $\frac{k-2}{n} \leq s \leq \frac{k+1}{n}$. If, for example, $\frac{k}{n} \leq s \leq \frac{k+1}{n}$, bound $|B(t) - B(s)|$ by

$$\left|B(t) - B\left(\frac{k-1}{n}\right)\right| + \left|B\left(\frac{k}{n}\right) - B\left(\frac{k-1}{n}\right)\right| + \left|B(s) - B\left(\frac{k}{n}\right)\right|,$$

which is at most the right side of (1.5).

Noting that the distribution of $Y_{k,n}$ does not depend on k, write for $\epsilon > 0$

$$P(\max_{1\leq k\leq nN} Y_{k,n} > \epsilon) \leq \sum_{k=1}^{nN} P(Y_{k,n} > \epsilon) = nNP(Y_{1,n} > \epsilon). \tag{1.6}$$

To check that the right side above tends to 0 as $n \to \infty$, we will apply Doob's inequality for discrete time submartingales – see Theorem A.33. If $0 < t_1 < \cdots < t_m$ are rational, $(B(t_1), ..., B(t_m))$ is a martingale, since the successive differences are independent and have mean 0. Therefore, $(B^4(t_1), ..., B^4(t_m))$ is a (nonnegative) submartingale. Doob's inequality gives

$$P(\max_{1\leq k\leq m} |B(t_k)| > \epsilon) \leq \frac{1}{\epsilon^4} EB^4(t_m).$$

Note that the bound on the right side depends on t_m, but not on m – this is very important in the next step. Applying this to a sequence of subsets that exhausts $Q \cap [0, \frac{1}{n}]$, and using the fact that $EB^4(t)$ is increasing in t, we see that

$$P(Y_{1,n} > \epsilon) \leq \frac{1}{\epsilon^4} EB^4\left(\frac{1}{n}\right).$$

Since $B(t)$ has the same distribution as $\sqrt{t}B(1)$ (the $N(0,t)$ distribution), we conclude that

$$P(Y_{1,n} > \epsilon) \leq \frac{EB^4(1)}{n^2\epsilon^4},$$

so that the right side of (1.6) tends to 0 as $n \to \infty$. Therefore, $\Delta_n \to 0$ in probability by (1.5) as required. □

Exercise 1.21. Use Doob's inequality, Theorem A.33, applied to even powers larger than the fourth in the proof of Theorem 1.20 to obtain the following improvement: for every $N > 0$ and every $0 < \alpha < \frac{1}{2}$ there is a random variable C so that

$$|B(t) - B(s)| \leq C|t - s|^{\alpha}, \quad 0 \leq s, t \leq N.$$

(We will see in Exercise 1.28 that this is not true if $\alpha = \frac{1}{2}$).

The following exercises develop another approach to the construction problem.

Exercise 1.22. Let $\{\phi_n\}$ be a complete orthonormal family in $L_2[0, 1]$, and define

$$\psi_n(t) = \int_0^t \phi_n(s)ds = \langle \phi_n, 1_{[0,t]} \rangle,$$

where $\langle \cdot \rangle$ is the usual inner product in $L_2[0, 1]$. Let $\{\xi_n\}$ be i.i.d. standard normal random variables. Use Corollary A.14 to show that for each $t \in [0, 1]$, the series

$$B(t) = \sum_n \xi_n \psi_n(t) \tag{1.7}$$

converges a.s. and in L_2, and that the resulting process B satisfies properties (a) and (b) in Proposition 1.11.

Exercise 1.23. Show that if $\{\xi_n\}$ are i.i.d. standard normal random variables, then for every $\epsilon > 0$,

$$\lim_{n \to \infty} \xi_n n^{-\epsilon} = 0 \quad a.s.$$

Exercise 1.24. In Exercise 1.22, take the orthonormal family to be the Haar functions that are defined as follows: $\phi_0 \equiv 1$,

$$\phi_{0,1}(t) = \begin{cases} +1 & \text{if } 0 \leq t \leq \frac{1}{2}; \\ -1 & \text{if } \frac{1}{2} < t \leq 1, \end{cases}$$

$$\phi_{1,1}(t) = \begin{cases} +\sqrt{2} & \text{if } 0 \leq t \leq \frac{1}{4}; \\ -\sqrt{2} & \text{if } \frac{1}{4} < t \leq \frac{1}{2}; \\ 0 & \text{if } \frac{1}{2} < t \leq 1, \end{cases} \qquad \phi_{1,2}(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq \frac{1}{2}; \\ +\sqrt{2} & \text{if } \frac{1}{2} < t \leq \frac{3}{4}; \\ -\sqrt{2} & \text{if } \frac{3}{4} < t \leq 1, \end{cases}$$

and in general, for $1 \leq k \leq 2^n$,

$$\phi_{n,k} = \begin{cases} +2^{n/2} & \text{if } \frac{k-1}{2^n} \leq t \leq \frac{k-(1/2)}{2^n}; \\ -2^{n/2} & \text{if } \frac{k-(1/2)}{2^n} < t \leq \frac{k}{2^n}; \\ 0 & \text{otherwise.} \end{cases}$$

Use Exercise 1.23 to show that the series in (1.7) converges uniformly on $[0, 1]$, and hence defines a standard Brownian motion.

The Brownian motion property is preserved by several transformations, as we now check.

Theorem 1.25. *Suppose that B is a standard Brownian motion. Then the following processes are also:*

$$X_1(t) = B(t+s) - B(s), \quad s > 0 \text{ fixed.}$$
$$X_2(t) = \frac{B(ct)}{\sqrt{c}}, \quad c > 0 \text{ fixed.}$$
$$X_3(t) = \begin{cases} tB(1/t) & \text{if } t > 0; \\ 0 & \text{if } t = 0. \end{cases}$$

Proof. The first two cases are left as an exercise. For the third, note that X_3 is a mean zero Gaussian process that is continuous except possibly at 0. To check the covariance, write

$$EX_3(s)X_3(t) = st \min(1/s, 1/t) = s \wedge t.$$

To check continuity at 0, write

$$\{\omega : \lim_{t \downarrow 0} B(t) = 0\} = \cap_{m \geq 1} \cup_{n \geq 1} \{\omega : |B(t)| \leq 1/m \text{ for all } t \in Q \cap (0, 1/n)\}$$

and

$$\{\omega : \lim_{t \downarrow 0} X_3(t) = 0\} = \cap_{m \geq 1} \cup_{n \geq 1} \{\omega : |X_3(t)| \leq 1/m \text{ for all } t \in Q \cap (0, 1/n)\}.$$

The right sides of these two identities have the same probability, since the two processes have the same finite dimensional distributions, and the events depend on the processes at only countably many times. The left side of the first has probability 1, since B has continuous paths. Therefore the left side of the second also has probability 1. □

Exercise 1.26. Check that $X_1(t)$ and $X_2(t)$ in Theorem 1.25 are Brownian motions.

Note that $B(n)$ is a sum of the i.i.d. increments $B(k) - B(k-1)$, so that $B(n)/n \to 0$ a.s. as $n \to \infty$ by Theorem A.15. The next two exercises show that Theorem 1.25 can be quite useful, in spite of its simplicity.

Exercise 1.27. Use the fact that $X_3(t)$ from Theorem 1.25 is a Brownian motion to show that $B(t)/t \to 0$ a.s. as $t \to \infty$.

Exercise 1.28. Use the fact that $X_3(t)$ from Theorem 1.25 is a Brownian motion and Corollary A.24 to show that

$$\limsup_{t \downarrow 0} \frac{B(t)}{\sqrt{t}} = +\infty \quad \text{and} \quad \liminf_{t \downarrow 0} \frac{B(t)}{\sqrt{t}} = -\infty \quad a.s.$$

In particular,

$$P\big(\forall \epsilon > 0, B(t) \text{ takes both signs in } [0,\epsilon)\big) = 1.$$

The next three exercises are designed to show that the property of path continuity is rather rare among continuous time processes. They also serve to introduce the symmetric stable processes (see Definition A.25 for the definition of a stable law) and compound Poisson processes. Note that Brownian motion is a symmetric stable process with $\alpha = 2$. The acronym "cadlag" (continue à droite, limites à gauche) refers to paths that are right continuous and have left limits. The cadlag property helps to resolve many measurability issues. For example, if $X(t)$ has cadlag paths, then the set

$$C = \{\omega : X(t,\omega) \text{ is continuous for } t \in [0,N]\}$$

is measurable, since it can be written as

$$C = \left\{\omega : \lim_{n\to\infty} \sup_{\substack{s,t\in Q\cap[0,N]\\ |s-t|\le\frac{1}{n}}} |X(s) - X(t)| = 0\right\}.$$

Exercise 1.29. Suppose that $(X(t), t \ge 0)$ is a stochastic process with the following three properties:

(i) It has stationary independent increments.

(ii) $X(t)$ has a symmetric stable law of index $\alpha \in (0,2]$. Specifically, take its characteristic function to be $Ee^{iuX(t)} = e^{-t|u|^\alpha}$. A consequence if $\alpha < 2$ is that $P(|X(1)| \ge x) \sim cx^{-\alpha}$ for some constant $c > 0$ as $x \to \infty$. (Note that this is false if $\alpha = 2$.)

(iii) Its sample paths are a.s. cadlag. (That there is a version of the process with this property will be proved in Chapter 3.)

(a) Prove that for each t, $X(\cdot)$ is a.s. continuous at t.

(b) Express

$$P\left(\max_{1\le k\le n}\left|X\left(\frac{k}{n}\right) - X\left(\frac{k-1}{n}\right)\right| \ge \epsilon\right)$$

in terms of the distribution of $X(1)$.

(c) Prove that if $\alpha < 2$, then

$$P(X(\cdot) \text{ is continuous on } [0,1]) = 0.$$

Exercise 1.30. Let $N(t)$ be a rate λ Poisson process, and Y_j be i.i.d. random variables independent of N. Define a continuous time process $(Z(t), 0 \le t \le 1)$ by

$$Z(t) = \sum_{j=1}^{N(t)} Y_j.$$

$Z(t)$ is known as a compound Poisson process.

(a) Compute the characteristic function of $Z(t)$ in terms of the characteristic function of Y_j.

(b) Show that Z has stationary independent increments.

(c) Prove that for each t, $Z(\cdot)$ is a.s. continuous at t.

(d) Compute

$$P(Z(\cdot) \text{ is continuous on } [0,t]).$$

Exercise 1.31. Let $Z_n(t)$ be as in the previous exercise with $\lambda = n$ and Y_j with characteristic function $\phi_n(u)$, where

$$\lim_{n\to\infty} n[1-\phi_n(u)] = |u|^\alpha,$$

with $0 < \alpha \le 2$. Show that the finite dimensional distributions of Z_n converge to those of the stable process from Exercise 1.29.

The following measurability fact will often be useful. For its statement, we assume that Ω has been modified so that $B(\cdot,\omega)$ is continuous for all ω.

Proposition 1.32. *$B(t,\omega)$ is jointly measurable in (t,ω).*

Proof. Let

$$X_n(t) = B\left(\frac{[nt]}{n}\right),$$

where $[\cdot]$ is the greatest integer function. This is jointly measurable for each n since for any Borel set A,

$$\{(t,\omega) : X_n(t,\omega) \in A\} = \bigcup_{k=0}^{\infty} \left[\frac{k}{n}, \frac{k+1}{n}\right) \times \left\{\omega : B\left(\frac{k}{n},\omega\right) \in A\right\},$$

and each set on the right is a measurable rectangle. By path continuity,

$$X_n(t,\omega) \to B(t,\omega)$$

uniformly on compact t sets. Therefore B is jointly measurable. □

1.6. Path properties

We noted one path property in Exercise 1.21 – Brownian paths are not only continuous, but are in fact locally Hölder continuous for any index $< \frac{1}{2}$. In Exercise 1.28, we saw that this is not the case if the index is $\frac{1}{2}$. We begin with a simple non-differentiability statement. Later, we will improve it. We use m below to denote Lebesgue measure.

Proposition 1.33. *(a) For every $t \ge 0$,*

$$P(B(\cdot,\omega) \text{ is not differentiable at } t) = 1.$$

(b) $m\{t \ge 0 : B(\cdot,\omega) \text{ is differentiable at } t\} = 0$ a.s.