BROWNIAN MOTION AND THE STRONG MARKOV PROPERTY

JAMES LEINER

Abstract. This paper is an introduction to Brownian motion. After a brief introduction to measure-theoretic probability, we begin by constructing Brownian motion over the dyadic rationals and extending this construction to $\mathbb{R}^d$. After establishing some relevant features, we introduce the strong Markov property and its applications. We then use these tools to demonstrate the existence of various Markov processes embedded within Brownian motion.

Contents

1. Introduction

With a varied array of uses across pure and applied mathematics, Brownian motion is one of the most widely studied stochastic processes. This paper seeks to provide a rigorous introduction to the topic, using [3] and [4] as our primary references.

In order to properly ground our discussion, however, we must first become comfortable with the basic framework of measure-theoretic probability. In this section, we provide a terse review of the subject sufficient to prepare the reader for some of the more advanced proofs within this paper. Most of the important results in this section are presented without proof and individuals wishing for a more detailed exposition of the topic can consult [1]. If so inclined, a large portion of the first half of the paper can be understood without measure theory. In this case, consult [2] for a more basic introduction to probability not requiring measure theory.

Definition 1.1. A **σ-algebra** on a set S is a subset of the power set $\Sigma \subseteq 2^S$ such that

(1) $S \in \Sigma$;
(2) if $A \in \Sigma$, then $A^c \in \Sigma$;

(3) if $A_1, A_2, ... \in \Sigma$, then $\bigcup_{i=1}^{\infty} A_i \in \Sigma$.

Note that by De Morgan's laws,

$$(\bigcup_{i=1}^{\infty} A_i^c)^c = \bigcap_{i=1}^{\infty} A_i.$$

Thus, a σ-algebra is closed under countable intersection as well.

Definition 1.2. A pair (S, Σ) where S is a set, and Σ is some σ-algebra over that set, is called a **measurable space**.

Definition 1.3. Let $\mathcal{C}$ be a collection of subsets of a set S. The σ-algebra generated by $\mathcal{C}$, denoted $\sigma(\mathcal{C})$, is the smallest σ-algebra on S such that $\mathcal{C} \subseteq S$. In other words, it is the intersection of all σ-algebras on S which have $\mathcal{C}$ as a subcollection.

Definition 1.4. Let S be a topological space. $\mathcal{B}(S)$, the **Borel σ-algebra**, is the σ-algebra generated by the family of open subsets of S.

Note that in the specific case of $\mathbb{R}$, we can say that

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(-\infty, x] : x \in \mathbb{R}\}).$$

Definition 1.5. Let (S, Σ) be a measurable space. A map $\mu : \Sigma \to [0, 1]$ is called a **measure** when $\mu(\{\varnothing\}) = 0$ and it is countably additive. That is,

$$\mu\left(\bigcup_{j=1}^{\infty} F_j\right) = \sum_{j=1}^{\infty} \mu(F_j).$$

A triple (Ω, Σ, μ) is then called a **measure space**.

Definition 1.6. Let (S, Σ) be a measurable space and let X be a function from Ω to $\mathbb{R}$. Then, X is Σ-measurable if $X^{-1}(H) \subseteq \Sigma$ for all $H \in \sigma(\mathbb{R})$.

Definition 1.7. Let (S, Σ, μ) be a measure space. When $\mu(\Sigma)$ equals 1, this map is termed a **probability measure** and the associated measure space is called a **probability space**.

We are now able to use this machinery to re-introduce some familiar concepts within probability theory. First, let us introduce some standard terminology. Conventionally, the measure space (S, Σ, μ) is denoted as $(\Omega, \mathcal{F}, \mathbb{P})$ and called a **probability triple**. The set Ω is called a **sample space**, while an element w of Ω is correspondingly termed a **sample point**. Similarly, the σ-algebra $\mathcal{F}$ is called a **family of events** and a particular element of $\mathcal{F}$ is termed an **event**.

Definition 1.8. Let $E_1, E_2, ..., E_n$ be a collection of events. The limit superior of these sets is defined as

$$\limsup E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n.$$

Because inclusion in the limit superior implies that $x \in E_n$ for infinitely many values, we henceforth denote $\limsup E_n$ as $E_n, i.o.$

Definition 1.9. A statement S about outcomes is said to be true **almost surely**, or **with probability one**, if $F = \{w : S(w) \text{ is true}\} \in \mathcal{F}$ and $\mathbb{P}(F) = 1$.

Lemma 1.10 (Borel-Cantelli Lemma)**.** *Suppose $E_1, E_2..., E_n$ is a collection of events such that*

$$\sum_{n=1}^{\infty} \mathbb{P}\{E_n\} < \infty.$$

Then, $\mathbb{P}(E_n, i.o.) = 0$.

Proof. Given a collection of events $\{E_n : n \in \mathbb{N}.\}$, we have for all k,

$$\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n \subset \bigcup_{n=k}^{\infty} E_n.$$

So for all k,

$$\mathbb{P}(E_n, i.o.) \leq \mathbb{P}(\bigcup_{n=k}^{\infty} E_n) \leq \sum_{n=k}^{\infty} \mathbb{P}(E_k).$$

However, by the hypothesis that $\sum_{n=1}^{\infty} \mathbb{P}\{E_n\}$ is finite, the sum on the right side of the equation converges to zero as k gets arbitrarily large. □

Definition 1.11. A **random variable** on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a map $X : \Omega \to \mathbb{R}$ that is $\mathcal{F}$-measurable.

We now turn to formalize the notion of independence and dependence with respect to events and random variables.

Definition 1.12. Let $\mathcal{F}$ be a σ-algebra. Sub-σ-algebras $\mathcal{G}_1, \mathcal{G}_2, ...$ of $\mathcal{F}$ are **independent** if whenever $G_i \in \mathcal{G}$ for all $i \in \mathbb{N}$ and $i_1, ...i_n$ are distinct, then

$$\mathbb{P}(G_{i_1} \bigcap G_{i_2} \bigcap ... \bigcap G_{i_n}) = \prod_{k=1}^{n} \mathbb{P}(G_{i_k}).$$

Definition 1.13. Random variables $X_1, X_2, ...$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are called independent if the σ-algebras $\sigma(X_1), \sigma(X_2), ...$ are independent.

Definition 1.14. Events $E_1, E_2, ...$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are called independent if the corresponding σ-algebras $\varepsilon_1, \varepsilon_2, ...$ are independent and $\varepsilon_i = \{\varnothing, E_n, \Omega \backslash E_n, \Omega\}$.

At this point, we will define the concept of an expectation formally. Doing so, however, first requires an understanding of what taking an integral with respect to a measure means exactly. It will be beyond the scope of this paper to develop this theory in sufficient detail. Interested readers can consult [1] for such a development.

Definition 1.15. Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The expectation of X, $E[X]$, is defined by

$$E[X] = \int_{\Omega} X d\mathbb{P} = \int_{\Omega} X(w)\mathbb{P}(dw).$$

Definition 1.16. Let X and Y be two random variables defined on a probability space. We define the covariance of X and Y, $Cov[X, Y]$, as

$$Cov[X, Y] = E[(X - E[X])(Y - E[Y])].$$

Definition 1.17. Let X be a random variable defined on a probability space. The variance of X, $Var[X]$, is defined as

$$Var[X] = Cov[X, X] = E[(X - E[X])^2].$$

Definition 1.18. Let X be a random variable defined on a probability space. We then define the **law** of X, $\mathcal{L}_x$, as

$$\mathcal{L}_x = \mathbb{P} \circ X^{-1},\ \mathcal{L}_x : \mathcal{B} \to [0, 1].$$

Thus, $\mathcal{L}_x$ defines a probability measure on $(R, \mathcal{B})$. We are now able to uniquely generate $\mathcal{L}_x$ by the function $F_x : \mathbb{R} \to [0, 1]$ defined as follows:

Definition 1.19. Let X be a random variable defined on a probability space. The **distribution function** of X is defined as

$$F_x(c) = \mathcal{L}_x(-\infty, c] = \mathbb{P}(X \leq c) = \mathbb{P}\{w : X(w) \leq c\}.$$

The properties of a distribution function remain as you would remember them from basic probability theory. In particular, we have the following:

(1) $F_x : \mathbb{R} \to [0, 1]$;
(2) F_x is increasing monotonically. That is, if $a \leq b$ then $F_x(a) \leq F_x(b)$;
(3) $\lim_{a\to\infty} F_x(a) = 1$ and $\lim_{a\to-\infty} F_x(a) = 0$;
(4) F is right-continuous.

Definition 1.20. Let X be a random variable defined on a probability space. We say that X has a **probability density function** f_X if there exists a function $f_X : X \to [0, \infty)$ such that

$$\mathbb{P}(X \in B) = \int_B f_X(x)dx,\ B \in \mathcal{B}.$$

In particular, we concern ourselves with a particular distribution function that is intimately tied with the concept of Brownian motion and will remain prominent over the course of this paper.

Definition 1.21. A random variable X has a **normal distribution** with mean μ and variance σ^2 if

$$\mathbb{P}\{X > x\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^\infty e^{\frac{-(u-\mu)^2}{2\sigma^2}} du.$$

Occasionally, the shorthand $N(\mu, \sigma^2)$ is used to refer to this distribution.

Within the course of this paper, we will also need to concern ourselves with continuous random variables in two dimensions. Although there is a natural way to extend the measure theory we have defined thus far to this purpose, it will be sufficient for our purposes to utilize the standard definitions from basic probability theory.

Definition 1.22. Let X and Y be random variables defined on a probability space. These variables are **jointly continuous** if there exists a **joint probability density function** f_{XY} such that for all $x, y \in \mathbb{R}$ and measurable sets of real numbers A and B, we have

$$\mathbb{P}(X \in A, Y \in B) = \int_B \int_A f_{XY}(x, y)dxdy.$$

Definition 1.23. Let X and Y be jointly continuous random variables defined on a probability space. We define the **conditional probability density function** as $f_Y(y \mid X = x)$ such that such that for all $x, y \in \mathbb{R}$ and every measurable set of real numbers B, we have

$$\mathbb{P}(Y \in B \mid X = x) = \int_B f_Y(y \mid X = x)dy = \int_B \frac{f_{XY}(x,y)}{f_X(x)}dy.$$

At this point, we are now ready to present the definition of Brownian motion.

Definition 1.24. A **stochastic process** is a collection of random variables $\{W_t : t \in \mathcal{T}\}$ on some probability space and indexed by some set of times $\mathcal{T}$.

Definition 1.25. A **Brownian motion** started at $x \in \mathbb{R}$ is a stochastic process with the following properties:

(1) $W_0 = x$;
(2) For every $0 \leq s \leq t$, $W_t - W_s$ has a normal distribution with mean zero and variance $t - s$, and $|W_t - W_s|$ is independent of $\{W_r : r \leq s\}$;
(3) With probability one, the function $t \to W_t$ is continuous.

A Brownian motion started at 0 is termed **standard Brownian motion**.

2. Construction

Given the definition of Brownian motion, one needs to present more work in order to demonstrate that the conditions imposed on the distributions allow for a continuous random process to exist. In this section, we present a construction of Brownian motion that demonstrates no contradictions arise.

Definition 2.1. The set of **non-negative dyadic rationals** is $\mathcal{D} = \bigcup_n \mathcal{D}_n$ where $\mathcal{D}_n = \{\frac{k}{2^n} : k = 0, 1, 2, ...\}$.

Definition 2.2. A standard one-dimensional Brownian motion on the dyadic rationals $\{W_q : q \in \mathcal{D}\}$ is a random process such that for every n, the random variables $W_{k/2^n} - W_{(k-1)/2^n}$, $k \in \mathbb{N}$ are independent and $N(0, \frac{1}{2^n})$.

Proposition 2.3. *Suppose X, Y are independent normal random variables, each $N(0,1)$. If*

$$Z = \frac{X}{\sqrt{2}} + \frac{Y}{\sqrt{2}} \text{ and } \hat{Z} = \frac{X}{\sqrt{2}} - \frac{Y}{\sqrt{2}},$$

then Z and $\widetilde{Z}$ are independent $N(0,1)$ variables.

Proof. We have

$$Z = \frac{1}{\sqrt{2}}N(0,1) + \frac{1}{\sqrt{2}}N(0,1) = N(0, (1/\sqrt{2})^2 + (1/\sqrt{2})^2) = N(0,1).$$

For convenience, let us define $\hat{X} = X/\sqrt{2}$ and $\hat{Y} = Y/\sqrt{2}$. We know these variables are each $N(0, 1/2)$ and that $Z = \hat{X} + \hat{Y}$. This allows us to write the joint density for $(\hat{X}, \hat{Y})$ as

$$\left(\frac{1}{\sqrt{2\pi(1/2)}}e^{-\hat{x}^2}\right)\left(\frac{1}{\sqrt{2\pi(1/2)}}e^{-\hat{y}^2}\right) = \frac{1}{\pi}e^{-(\hat{x}^2+\hat{y}^2)}.$$

The joint density for $(\hat{X}, Z)$ becomes $\frac{1}{\pi}e^{-(\hat{x}^2+(z-\hat{x})^2)}$ and the density for Z becomes $\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$. Using these, we can say that the conditional density of $\hat{X}$ given $Z = z$ is

$$\frac{(1/\pi)e^{-(\hat{x}^2+(z-\hat{x})^2)}}{(1/\sqrt{2\pi})e^{-z^2/2}} = \frac{1}{\sqrt{\pi/2}}e^{-2(\hat{x}-z/2)^2}.$$

So, conditioned on the value of Z, we know that $\hat{X}$ is $N(Z/2, 1/4)$. Similarly, we know that conditioned on the value of Z that $\hat{Y}$ is $N(Z/2, 1/4)$. This allows us to write

$$\hat{X} = \frac{Z}{2} + \frac{\widetilde{Z}}{2} \text{ and } \hat{Y} = \frac{Z}{2} - \frac{\widetilde{Z}}{2}$$

where $\widetilde{Z}$ is $N(0, 1)$ and independent of Z. Substituting back X and Y and rearranging terms yields the desired equations.

□

Lemma 2.4. *Standard Brownian motion exists over the dyadic rationals.*

Proof. Define $J(k, n)$ as follows

$$J(k, n) = 2^{n/2}[W_{k/2^n} - W_{(k-1)/2^n}].$$

We start with the assumption that there exist a countable number of independent normal random variables $\{Z_n \colon n \in \mathbb{N}\}$ and work back recursively to define W_q. We can start by defining $\{J(k, 0) = Z_k : k \in \mathbb{N}\}$. Now, assume that there exists n such that $\{J(k, n) : k \in \mathbb{N}\}$ has been defined only using $\{Z_q : q \in \mathbb{D}\}$ so that they are independent $N(0, 1)$ variables. We can then define $J(k, n + 1)$ as follows:

$$J(2k-1, n+1) = \frac{J(k,n)}{\sqrt{2}} + \frac{Z_{(2k+1)/2^{n+1}}}{\sqrt{2}}$$
$$J(2k, n+1) = \frac{J(k,n)}{\sqrt{2}} - \frac{Z_{(2k+1)/2^{n+1}}}{\sqrt{2}}$$

Using Proposition 2.3 repeatedly defines $\{J(k, n + 1) : k \in \mathbb{N}\}$, thus yielding a collection of independent $N(0, 1)$ variables. We are then able to construct $W_{k/2^n}$ in a natural way:

$$W_{k/2^n} = 2^{-n/2}\sum_{j=1}^{k} J(j, n)$$

So,

$$2^{n/2}(W_{k/2^n} - W_{(k-1)/2^n}) = \sum_{j=1}^{k} J(j, n) - \sum_{j=1}^{k-1} J(j, n) = J(k, n).$$

□

Lemma 2.5. *If W_q, $q \in \mathcal{D}$ is a standard one-dimensional Brownian motion, then almost surely, the function converges uniformly on every interval closed interval $[a, b]$.*

Proof. We begin by defining

$$M_n = sup\{|W_s - W_t| : 0 \leq s, t \leq 1, |s - t| \leq 2^{-n}, s, t \in \mathcal{D}\}.$$

First, note that the statement $M_n \to 0$ as $n \to 0$ is true if and only if the function $t \longmapsto W_t$ converges uniformly. If M_n converges to zero, then given ϵ we can choose $M_n < \epsilon$ and then pick $\delta = 2^{-n}$ to ensure that $|W_s - W_t| < \epsilon$ for all $|s-t| < \delta$. Similarly, given a uniformly continuous function then for any ϵ, we can set $2^{-n} < \delta$ which ensures that $M_n < \epsilon$.

Keeping this in mind, we define

$$K_n = \max_{k=1,..2^n} \sup \left\{ |W_q - W_{(k-1)/2^n}| : q \in \mathcal{D}, \frac{k-1}{2^n} \leq q \leq \frac{k}{2^n} \right\}.$$

It is clear that $K_n \leq M_n$ forms a bound because the dyadic rationals contained in an interval are a subset of the real numbers. We can then bound K_n above using the triangle inequality.

$$|W_s - W_t| \leq |W_s - W_{(k-1)/2^n}| + |W_t - W_{(k)/2^n}| + |W_{kn} - W_{(k-1)/2^n}| \leq 3M_n$$

Thus,

$$K_n \leq M_n \leq 3K_n.$$

We must now demonstrate that $K_n \to 0$. First, let us define the following:

$$\tilde{K} = \sup\{W_q : q \in \mathcal{D} \bigcap [0,1]\}$$

$$\tilde{L} = \max\{W_q : q \in \mathcal{D}_n \bigcap [0,1]\}$$

$$\tilde{M} = \max\{W_{k/2^n} : k = 1, ..., 2^n\}$$

For fixed n, define the event E_k as the first time that $\tilde{M}$ is greater than or equal to some constant a. More formally, let it be the event that

$$\{W_{k/2^n} \geq a, W_{j/2^n} \text{ for } j = 1, ..., k-1\}.$$

From this definition, it is clear that the events $E_1, ..., E_n$ are pairwise disjoint. That is, $E_k \bigcap E_j = \varnothing$ for $j \neq k$. Moreover, the union over this entire set of events is equal to the event that $\tilde{M} \geq a$. Lastly, we can tell that the random variable $W_1 - W_{k/2^n}$ is independent of the event E_k for all possible k. Therefore, we can write

$$\begin{aligned} \mathbb{P}[E_k \bigcap \{W_1 \geq a\}] &\geq \mathbb{P}[E_k \bigcap \{W_1 - W_{k/2^n} \geq 0\}] \\ &= \mathbb{P}(E_k)\mathbb{P}\{W_1 - W_{k/2^n} \geq 0\} \geq \frac{1}{2}\mathbb{P}(E_k). \end{aligned}$$

Thus,

$$\frac{1}{2}\mathbb{P}\{\tilde{M} \geq a\} = \frac{1}{2}\sum_{k=1}^{2^n} \mathbb{P}(E_k) \leq \sum_{k=1}^{2^n} \mathbb{P}[E_k \bigcap \{W_1 \geq a\}] = \mathbb{P}\{W_1 \geq a\}.$$

So,

$$\mathbb{P}\{\tilde{M} \geq a\} \leq 2\mathbb{P}\{W_1 \geq a\}.$$

Now note that if $\tilde{K} > a$, then $W_q > a$ for some $q \in \mathcal{D} \bigcap [0,1]$. This implies that

$$\mathbb{P}\{\tilde{K} > a\} \leq \lim_{n \to \infty} \mathbb{P}\{\tilde{L} \geq a\}.$$

So,

$$\mathbb{P}\{\tilde{K} > a\} \leq 2\mathbb{P}\{W_1 \geq a\}. \tag{2.6}$$

Now, let us remind ourselves that given any $a \in \mathcal{D}$, Brownian motion can be bounded as follows:

$$\int_a^\infty e^{-x^2/2}dx \leq \int_a^\infty e^{-a^2/2}dx = \frac{2}{a}e^{-a^2/2} \Rightarrow \mathbb{P}\{W_1 \geq a\} \leq \frac{1}{\sqrt{2\pi}}\int_a^\infty e^{-x^2/2}dx \leq \frac{1}{a}e^{\frac{-a^2}{2}}$$

This, together with Equation 2.6 yields

$$\mathbb{P}\{\tilde{K} > a\} \leq \frac{2}{a}e^{-a^2/2}.$$

Thus,

$$\sup\{|W_q| : q \in \mathcal{D}\bigcap[0,1]\} \leq \frac{4}{a}e^{-a^2/2}.$$

Finally, consider $2^{n/2}K_n$, which is merely the supremum of 2^n different variables all with the same distribution as $\sup\{|W_q| : q \in \mathcal{D}\bigcap[0,1]\}$. Given a fixed $a \in \mathbb{R}$, the probability that the supremum of a collection of independent random variables is greater than a is merely the sum of each individual probability. We can therefore write

$$\mathbb{P}\{K_n \geq a2^{-n/2}\} \leq 2^n \sup\{|W_q| : q \in \mathcal{D}\bigcap[0,1]\} \leq \frac{4}{a}e^{-a^2/2}.$$

Setting $a = 2\sqrt{n}$ yields

$$\mathbb{P}\{K_n \geq 2\sqrt{n}2^{-n/2}\} \leq \frac{2}{\sqrt{n}}(2/e^2)^n \Rightarrow \sum_{n=1}^{\infty}\mathbb{P}\{K_n \geq 2\sqrt{n}2^{-n/2}\} < \infty.$$

Appling the Borel-Cantelli lemma implies that with probability one, $K_n < 2\sqrt{n}2^{-n/2}$ for all n large enough. Thus, $K_n \to 0$ as $n \to \infty$, which completes the proof.

□

Theorem 2.7. *Standard Brownian motion exists.*

Proof. This follows naturally from uniform continuity. Let us choose $\frac{\epsilon}{2}$ which yields some δ such that $|W_t - W_s| \leq \frac{\epsilon}{2}$ for all $s, t \in \mathcal{D}$. Now pick $n_0 \in \mathbb{N}$ such that $\frac{1}{2^{n_0}} < \delta$.

Pick $a \in \mathcal{D}$. Now, pick $n, m > n_0$ and $k_{n_0} = 0, 1, ..., 2^{n_0}$ such $0 < a - \frac{k_{n_0}}{2^n} < \frac{1}{2^n}$. Picking k_n and k_m in a similar fashion yields the following:

$$|k_n - k_{n_0}| < \frac{1}{2^{n_0}}$$
$$|k_m - k_{n_0}| < \frac{1}{2^{n_0}}$$

So,

$$|W_{k_n} - W_{k_m}| \leq |W_{k_n} - W_{k_{n_0}}| + |W_{k_m} - W_{k_{n_0}}| < \epsilon.$$

This yields a Cauchy sequence W_{k_n} in $\mathbb{R}$, which in turn defines a convergent subsequence with a unique limit. Defining $\{W_t, t \in \mathbb{R}\}$ as this limit yields a unique extension of $\{W_q : q \in \mathbb{D}\}$ to $\{W_t : t \in \mathbb{R}\}$ that is continuous.

□

3. Non-Differentiability

Despite being continuous, the random nature of Brownian motion yields many interesting pathological properties. The most prominent example of this is that it is nowhere differentiable. We present a proof of this for the interval $[0,1]$ for convenience, but this can be easily generalized for any arbitrary closed interval $[a,b]$.

Theorem 3.1. *W_t is nowhere differentiable on $t \in [0,1]$ with probability one.*

Proof. Define $M(k,n)$ and M_n as follows:

$$M(k,n) = max\{|W_{\frac{k}{n}} - W_{\frac{k-1}{n}}|, |W_{\frac{k+1}{n}} - W_{\frac{k}{n}}|, |W_{\frac{k+2}{n}} - W_{\frac{k+1}{n}}|\}$$

$$M_n = min\{M(1,n), ..., M(n,n)\}$$

Suppose there exists a $t \in [0,1]$ at which W_t is differentiable. Then, for some $M \in \mathbb{R}$,

$$\lim_{r \to t} \frac{W_r - W_t}{t - t} = M.$$

So, for any $\epsilon > 0$, there exists δ such that $|W_r - W_t| < (M+\epsilon)|r-t|$ whenever $|r-t| < \delta$.

Given an ϵ, pick n_o such that $\frac{2}{n_o} < \delta$. Then, for all $n \geq n_o$, $\frac{2}{n} \leq \frac{2}{n_o} < \delta$.

This implies that for $\frac{k}{n}$ such that $0 \leq t - \frac{k}{n} < \frac{1}{n}$,

$$|W_{\frac{k}{n}} - W_{\frac{k-1}{n}}| \leq |W_t - W_{\frac{k-1}{n}}| + |W_t - W_{\frac{k}{n}}| \leq (M+\epsilon)(2/n) + (M+\epsilon)(2/n) \leq 4(M+\epsilon)/n.$$

Setting $C = 4(M+\epsilon)$ and using a similar argument to bound the other two terms demonstrates that if there exists a $t \in [0,\infty]$ at which W_t is differentiable, then there exists a $C, n_o < \infty$ such that for all $n \geq n_o$, $C/n \geq M_n$.

Using property (2) of the definition of Brownian motion, we can see that $M(k,n)$ takes the maximum of a collection of independent, normally distributed random variables all with mean zero and variance $\frac{1}{n}$. This implies that for any Brownian Motion and any k, n,

$$\begin{aligned} \mathbb{P}\{M(k,n) \leq C/m\} &\leq |P\{|W_{1/n}| \leq C/n\}|^3 \\ &\leq \left(\sqrt{\frac{2n}{\pi}} \int_0^{C/n} e^{\frac{-x^2}{2n}} dx \right)^3 \leq \left(C\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} \right)^3. \end{aligned}$$

So, there exists a c such that $\mathbb{P}\{M(k,n) \leq C/m\} \leq (\frac{c}{\sqrt{n}})^3$. Hence, this yields

$$\mathbb{P}\{M_n \leq C/n\} \leq \mathbb{P}\{\bigcup_{k=1}^{n} \{M(k,n) \leq C/n\}\} \leq \sum_{k=1}^{n} \mathbb{P}\{M(k,n) \leq C/n\} \leq \frac{c}{\sqrt{n}}.$$

Thus,

$$\lim_{n \to \infty} \mathbb{P}\{M_n \geq C/n\} = 1.$$

Thus, $M_n \geq C/n$ with probability one for arbitrarily large enough n. Suppose W_t is differentiable somewhere with positive probability. Then, with positive probability, $C/n \geq M_n$ for $n \geq n_0$. Thus, $M_n \leq C/n$ for arbitrarily large n. Thus, W_t is almost surely not differentiable on $t \in [0, 1]$. □

4. Invariance Properties

In this section, we begin to prove some basic properties of Brownian motions that will become invaluable as we start delving into more complex proofs.

Lemma 4.1 (Scaling invariance)**.** *Suppose W_t is a standard Brownian motion and let $a > 0$. Then, the process $X_t = \frac{1}{a} W_{a^2 t}$ is also a standard Brownian motion.*

Proof. Under scaling, continuity of paths and independence of increments still hold. Considering $W_t - Ws$, we have $\frac{1}{a^2} N(0, a^2(t-s))$. This has a mean of zero and a variance of $\frac{1}{a^2} a^2(t-s) = t - s$. □

Theorem 4.2 (Time inversion)**.** *Suppose W_t is a standard Brownian motion. Then, the process defined by X_t,*

$$X_t = \begin{cases} 0 & \text{if } t = 0 \\ tW_{1/t} & \text{if } t > 0 \end{cases}.$$

is a standard Brownian motion.

Proof. The increments of this process having an expected value of zero is immediate. To see that the other properties are satisfied, note that for Brownian motions, we have

$$\begin{aligned} Cov[W_t, W_{t+s}] = E[W_t W_{t+s}] &= E[W_t(W_{t+s} - W_t + W_t)] \\ &= E[W_t^2] + E[W_{t+s}]E[W_t - W_s] = t. \end{aligned}$$

Then, for X_t we get

$$\begin{aligned} Cov[X_t, X_{t+s}] &= Cov[tW_{1/t}, (t+s)W_{1/(t+s)}] \\ &= t(t+s)Cov[W_{1/t}, W_{1/(t+s}] = (t+s)\frac{t}{t+s} = t. \end{aligned}$$

Thus,

$$Cov[X_t, X_{t+s} - X_t] = Cov[X_t, X_{t+s}] - Var[X_t] = t - t = 0.$$

$$\begin{aligned} Var[X_{t+s} - X_t] &= Var[X_{t+s}] + Var[X_t] - 2Cov[X_{t+s}, X_t] \\ &= (t+s) + t - 2t = s. \end{aligned}$$

This shows that the increments have the right variance. Independence of increments holds due to X_t and X_{t+s} having zero covariance and being normal variables.

Lastly, we need to demonstrate continuity. When $t > 0$, this is clear. Now, recall that the distribution of X_t over the rationals is the same as the distribution for a Brownian Motion. This implies that for $t \in \mathbb{Q}$,

$$\lim_{t \to 0} X_t = 0.$$

Thus, completing the proof. □

Corollary 4.3 (Law of large numbers)**.** *Almost surely, $\lim_{t \to \infty} \frac{W_t}{t} = 0$.*

Proof. Let X_t be defined as in the above theoreom. Then, we have $\lim_{t\to\infty} \frac{W_t}{t} = \lim_{t\to\infty} X_{1/t} = 0$. □

5. Markov Properties

In this section, we begin to discuss multi-dimensional Brownian motion. Let us start with this definition.

Definition 5.1. If $W_1, ..., W_d$ are all independent Brownian motions started in $x_1, ..., x_d$, then the random process W_t given by

$$W_t = (W_1, ..., W_d).$$

is called a **d-dimensional Brownian motion** started in $(x_1, ..., x_d)$. If W_t starts at the origin it is termed a **standard d-dimensional Brownian motion**.

Definition 5.2. A filtration on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $\mathcal{F}(t) : t \geq 0)$ of σ-algebras such that $\mathcal{F}(S) \subset \mathcal{F}(t) \subset \mathcal{F}$. A probability space with a filtration is termed a filtered probability space. A random process $\{X_t : t \geq 0\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is adapted if X_t is $\mathcal{F}(t)$ measurable for all $t \geq 0$.

Keeping these definitions in mind, we begin by establishing the simple Markov Property. Intuitively, this states that knowing the current position of a Brownian motion yields as much information as knowing the entire history of positions up to that point.

Theorem 5.3 (Markov Property)**.** *Let $\{W_t : t \geq 0\}$ is a Brownian motion started in $x \in \mathbb{R}^d$. Then the process $\{W_{t+s} - W_s : t, s > 0\}$ is a Brownian motion started at the origin and is independent of $\{W_t : 0 \leq t \leq s\}$.*

Proof. This follows directly from property the independence of increments of Brownian motion. □

However, this is rather trivial. A preliminary means of making this property slightly stronger is establishing that Brownian motion is independent of information that exists an infinitesimal amount of time into the future.

Definition 5.4. The **germ σ-algebra** is defined as $\mathcal{F}^+(0)$, where

$$\mathcal{F}^+(t) = \bigcap_{s>t} \mathcal{F}^0(s)$$

and $\{\mathcal{F}^0 : t \geq 0\}$ is the σ-algebra generated by $\{W_t : 0 \leq s \leq t\}$.

Definition 5.5. The **tail σ-algebra**, $\mathcal{T}$ of a Brownian motion is defined as

$$\mathcal{T} = \bigcap_{t\geq 0} \mathcal{G}(t),$$

where $\mathcal{G}(t)$ is the σ algebra generated by $\{W_s : s \geq t\}$.

Theorem 5.6. *For all $s \geq 0$, the random process $\{W_{t+s} - W_s : t \geq 0\}$ is independent of $\mathcal{F}^+(s)$.*

Proof. By continuity, we can write the following for a strictly decreasing sequence $\{s_n : n \in \mathbb{N}\}$ converging to s:

$$W_{t+s} - W_s = \lim_{n\to\infty} W_{s_n+t} - W_{s_n}$$

However, the Markov property verifies that the right side of the above equation is independent of $\mathcal{F}^+(s)$. □

Definition 5.7. We define $\mathbb{P}_x$ as the probability measure which makes the random process $\{W_t : t \geq 0\}$ a d-dimensional Brownian motion started in $x \in \mathbb{R}^d$.

Theorem 5.8 (Blumenthal's 0-1 law)**.** *Let $x \in \mathbb{R}^d$ and $A \in \mathcal{F}^+(0)$. Then $\mathbb{P}_x(A) \in \{0, 1\}$.*

Proof. By Theorem 5.11, we know that any $A \in \sigma\{W_t : t \geq 0\}$ is independent of $\mathcal{F}^+(0)$. However, because $\sigma\{W_t : t \geq 0\} \subset \mathcal{F}^+(0)$ means that any A chosen from $\mathcal{F}^+(0)$ must be independent from itself. This can only be the case if it has probability zero or one. □

Theorem 5.9 (Kolgomorov's 0-1 law)**.** *Let $x \in \mathbb{R}^d$ and $A \in \mathcal{T}$. Then $\mathbb{P}_x(A) \in \{0, 1\}$.*

Proof. Let $A \in \mathcal{T}$. We can map the tail σ-algebra onto the germ σ-algebra as follows. First, let W_t be a Brownian motion and define $X_{1/s} = \frac{1}{s}W_s$. Now, define $S(t) = \sigma(W_s : s \geq t)$. We then have

$$S(t) = \sigma(sX_{1/s} : s \geq t) = \sigma(X_{1/s} : s \geq t) = \sigma(X_u : u \leq 1/t).$$

Thus,

$$\mathcal{T} = \bigcap_{t \geq 0} S(t) = \bigcap_{t \geq 0} \sigma(X_u : u \leq 1/t) = \mathcal{F}^+(0).$$

Then for any $x \in \mathbb{R}^d$, we know $\mathbb{P}_x(A)$ is either zero or one by Blumenthal's 0-1 law. □

Definition 5.10. A random variable T with values in $[0, \infty)$ defined on a filtered probability space is called a stopping time if $\{T < t\} \in \mathcal{F}(t)$ for every $t \geq 0$. It is called a strict stopping time if for every $t \geq 0$, $\{T \leq t\} \in \mathcal{F}(t)$.

Remark 5.11. Every strict stopping time is also a stopping time.

Theorem 5.12. *Every stopping time with respect to the filtration $\{\mathcal{F}^+(t) : t \geq 0\}$ is a strict stopping time.*

Proof. First, let us establish the right-continuity of $\{\mathcal{F}^+(t) : t \geq 0\}$. To do this, we can write

$$\mathcal{F}^+(t) = \bigcap_{n=1}^{\infty} \bigcap_{k=1}^{\infty} \mathcal{F}^0(t + \frac{1}{n} + \frac{1}{k}) = \bigcap_{\epsilon > 0} \mathcal{F}^+(t + \epsilon).$$

Thus,

$$\{T \leq t\} = \bigcap_{k=1}^{\infty} \{T < t + \frac{1}{k}\} \in \bigcap_{n=1}^{\infty} \mathcal{F}^+(t + \frac{1}{n}) = \mathcal{F}^+(t).$$

□

Theorem 5.13 (Strong Markov property)**.** *For every almost surely finite stopping time T, the process $\{W_{T+t} - W_T : t \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}^+(T)$.*

Proof. Let T be a stopping time. We can then define

$$T_n = (m+1)2^{-n}, \text{ where } m/2^n \leq T < (m+1)/2^n.$$

This can be thought of as a discrete approximation which stops at the first dyadic rational next to the original. Keeping in mind that this definition implies that T_n is also as stopping time, we then define the following:

$$W_k(t) = W_{t+k/2^n} - W_{k/2^n} \text{ and } W_k = \{W_k(t) : t \geq 0\}$$

$$W_*(t) = W_{t+T_n} - W_{T_n} \text{ and } W_* = \{W_*(t) : t \geq 0\}$$

Now, take $E \in \mathcal{F}^+(T_n)$ and the event $\{W_* \in A\}$. We have

$$\mathbb{P}(\{W_* \in A\} \bigcap E) = \sum_{k=1}^{\infty} \mathbb{P}(\{W_k \in A\} \bigcap E \bigcap \{T_n = k/2^n\}).$$

Note, however, that $E \bigcap \{T_n = k/2^n\} \in \mathcal{F}^+(k/2^n)$, which by Theorem 5.12 is independent of $\{W_k \in A\}$. Thus, we have

$$\mathbb{P}(\{W_* \in A\} \bigcap E) = \sum_{k=0}^{\infty} \mathbb{P}\{W_k \in A\}\mathbb{P}(E \bigcap \{T_n = k/2^n\}).$$

Now, using the Markov property we see that for all $k \in \mathbb{N}$, $\mathbb{P}\{W \in A\} = \mathbb{P}\{W_k \in A\}$. This yields

$$\begin{aligned} \sum_{k=0}^{\infty} \mathbb{P}\{W_k \in A\}\mathbb{P}(E \bigcap \{T_n = k/2^n\}) &= \mathbb{P}\{B \in A\} \sum_{k=0}^{\infty} \mathbb{P}(E \bigcap \{T_n = k/2^n\}) \\ &= \mathbb{P}\{B \in A\}\mathbb{P}(E). \end{aligned}$$

Thus, W_* is independent of every E and hence independent of $\mathcal{F}^+(T_n)$. Now, recall that the sequence T_n is a uniformly decreasing sequence that converges to T, hence $\mathcal{F}^+(T_n) \subset \mathcal{F}^+(T)$ is independent of the Brownian motion $W_{s+T_n} - W_{T_n}$. Then, the random process $W_{r+T} - W_T$, defined by the increments

$$W_{s+t+T} - W_{t+T} = \lim_{n \to \infty} W_{s+t+T_n} - W_{t+T_n},$$

is independent, $N(0, s)$, and almost surely continuous. Thus, it is a Brownian motion and independent of $\mathcal{F}^+(T)$. □

6. Applications

Theorem 6.1 (Reflection principle)**.** *If T is a stopping time and $\{W_t : t \geq 0\}$ is a standard Brownian motion, then the random process*

$$W_t^* = \begin{cases} 2W_T - W_t & \text{if } t \leq T \\ W_t & \text{if } 0 \leq t \leq T \end{cases}$$

is a standard Brownian motion.

Proof. Consider the following random processes:

$$\{W_{t+T} - W_T : t \geq 0\} \tag{6.2}$$

$$\{-(W_{t+T} - W_T) : t \geq 0\} \tag{6.3}$$

Using the Strong Markov property, we know that both (6.2)and (6.3) are Brownian motions and independent of the beginning process $\{W_t : 0 \geq t \geq T\}$. Consider that both (6.2) and (6.3) begin at the end point of this beginning process. So, if we attach (6.2) to the end point of the beginning process, we merely get the continuous process $\{W_t : t \geq 0\}$. If we attach (6.3) to the beginning process, then for all $t \geq T$, we have $W_T - (W_t - W_T) = 2W_T - W_t$. This gives $\{W_t^* : t \geq 0\}$. Because (6.2) and (6.3) have the same distribution and are Brownian motion, we therefore know that both of these concatenated processes have the same distributions and are Brownian motions. □

We now focus on the properties of the zero set of Brownian motion.

Theorem 6.4. *Let $\{W_t : t \geq 0\}$ be a Brownian motion. Then, the zero set of Brownian motion*

$$\textbf{Zeros} = \{t \geq 0 : W_t = 0\}$$

is a closed set with no isolated points.

Proof. Note that $\{0\}$ is a closed set. The zero set being closed is then immediate from the continuity of Brownian motion, as the preimage of a closed set has to be closed. Now, consider the construction $\tau_q = \inf\{t \geq q : W_t = 0\}$ for all rationals $q \in [0, \infty)$. First, note that the infimum of this set is also a minimum because it is a closed set. Second, note this construction is an almost surely finite stopping time. Now, apply the strong Markov property to τ_q. This yields a new Brownian motion started at 0, $W_{T+\tau_q} - W_{\tau_q}$. However, we know that any infinitesimally small interval to the right of the origin will contain a zero. Thus, almost surely τ_q is not isolated from the right for all positive rationals.

Now, consider the remaining points in the Zero set that do not correspond to some τ_q. Fix such a point, t. We can then pick a sequence of rational numbers $q(n)$ converging to t. Then, the sequence $\tau_q(n)$ converges to t. Thus, t is not isolated from the left. □

An interesting corollary to this is that the zero set is uncountable.

Theorem 6.5. *A closed set with no isolated points is uncountable.*

Proof. Let A be a closed set with no isolated points. Since A is a collection of limit points, we know that it cannot be finite. Thus, A can either be countably or uncountably infinite. Suppose that it is countable. We can then write the set as $A = \{a_1, a_2, a_3, ..., a_n\}$. Define $U_n := (a_n - 1, a_n + 1)$. Then, for every n we can construct sets such that

(1) $\bar{U}_{n+1} \subseteq U_n$;
(2) U_n does not contain any points a_j for $j < n$;
(3) U_n contains a_n.

Then, consider the set $V = \bigcap_{n\in\mathbb{N}} \bar{U}_n \bigcap A$. Each of the sets $\bar{U_n} \bigcap A$ is compact. Because an intersection of nested compact sets is non-empty, we therefore know that V contains a point not enumerated in $\{a_1, a_2, a_3, ..., a_n\}$. This yields a contradiction. □

7. Derivative Markov Processes

Using the tools that we have built up to this point, we can now show the existence of several important Markov processes embedded within Brownian motion in interesting ways.

Definition 7.1. A function $p : [0, \infty) \times \mathbb{R}^d \times \mathcal{B} \to \mathbb{R}$, where $\mathcal{B}$ denotes the Borel sigma algebra over $\mathbb{R}^d$ is a **Markov transition kernel** provided that the following hold:

(1) $p(\cdot, \cdot, A)$ is measurable as a function of (t,x) for each $A \in \mathcal{B}$;
(2) $p(t, x, \cdot)$ is a Borel probability measure on $\mathbb{R}^d$ for all $t \geq 0$ and $x \in \mathbb{R}^d$; When integrating a function f with respect to this probability measure, we write

$$\int f(y)p(t, x, dy);$$

(3) For all $A \in \mathbb{R}$, $x \in \mathbb{R}^d$ and $t, s > 0$,

$$p(t + s, x, A) = \int_{\mathbb{R}^d} p(t, y, A)p(s, x, dy).$$

.

Theorem 7.2. *For any $a \geq 0$, define the stopping times*

$$T_a = \inf\{t \geq 0 : W_t = a\}.$$

Then, $\{T_a : a \geq 0\}$ is an increasing Markov process with transition kernel given by the densities

$$p(a, t, s) = \frac{a}{\sqrt{2\pi(s-t)^3}} e^{-\frac{a^2}{2(s-t)}} 1\{s \geq t\}. \tag{7.3}$$

This is called the ***stable suboordinator of index 1/2****.*

Proof. Fix a, b such that $0 \leq b \leq a$. Now, note that for all $t \geq 0$, we have

$$\{T_a - T_b = t\} = \{W_{T_b+s} - W_{T_b} < a - b, \text{ for } s < t, \text{ and } W_{T_b+t} - W_{T_b}\}.$$

Using the strong Markov property, we know that this event is independent of $\mathcal{F}^+(T_b)$. We therefore can establish independence with $\{T_d : d \leq b\}$. This, in turn, establishes the Markov property of $\{T_a : a \geq 0\}$. Now, we are able to determine the transition Markov kernel through the reflection principle. We have

$$\begin{aligned}
\mathbb{P}\{T_a - T_b \leq t\} &= \mathbb{P}\{T_{a-b} \leq t\} \\
&= \mathbb{P}\{\max_{o \leq s \leq t} W_s \geq a - b\} = 2\mathbb{P}\{W_t \geq a - b\} \\
&= \int_{a-b}^{\infty} \frac{1}{\sqrt{t\pi s^3}} (a-b) e^{-\frac{(a-b)^2}{2s}} ds \\
&= \int_0^t \frac{1}{\sqrt{2\pi s^3}} (a-b) e^{-\frac{(a-b)^2}{2s}} ds.
\end{aligned}$$

Using the substitution $\sqrt{t/s}(a - b)$ yields the new integral

$$\int_0^t \frac{a}{\sqrt{2\pi(s-t)^3}} e^{-\frac{a^2}{2(s-t)}} ds.$$

□

Analogously, we can find a similar sort of process embedded within two dimensional Brownian motion.

Theorem 7.4. *Let* $\{W_t : t \geq 0\} = (W_1(t), W_2(t))$ *be a two-dimensional Brownian motion. Given a fixed* $a \geq 0$*, we then define*

$$V(a) = \{(x, y) \in \mathbb{R}^2 : x = a\}$$

and let $T(a)$ *be the first hitting time of* $V(a)$*. Defining a new process* $\{X(a) : a \geq 0)\}$ *such that* $X(a) = W_2(T(a))$ *then yields a Markov process with transition kernel given by*

$$p(a, x, A) = \frac{1}{\pi} \int_A \frac{a}{a^2 + (x - y)^2} dy. \tag{7.5}$$

This is termed the ***Cauchy process****.*

Proof. First, note that $T(a) < T(b)$ for all $a < b$. This, together with the strong Markov property of Brownian motion applied to the stopping time $T(a)$, confirms that the Markov property applies to $X(a)$.

Now, recall that Theorem 7.2 confirms that $T(a)$ has the density given by (7.3). This, coupled with the independence of $T(a)$ from $W_2(s)$ allows us to write the density of $W_2(T(a))$ as

$$\int_0^\infty \frac{1}{\sqrt{2\pi s}} e^{-\frac{x^2}{2s}} \frac{a}{\sqrt{2\pi s^3}} e^{-\frac{a^2}{2s}} ds.$$

Substituting $\sigma = \frac{1}{2s}(a^2 + x^2)$ then yields

$$\int_0^\infty \frac{ae^{-\sigma}}{\pi(a^2 + x^2)} d\sigma = \frac{a}{\pi(a^2 + x^2)}.$$

□

Acknowledgments. It is a pleasure to thank my mentor, Marcelo Alvisio for his helpful critiques on the composition of this paper, as well as his help in building up the requisite level of mathematical maturity needed to explore this fascinating topic. I would also like to voice my appreciation for Peter May and the rest of the faculty who gave up their time to organize and teach in the program. Without their help, this paper would not have been possible.

References

[1] David Williams. Probability with Martingales. Cambridge University Press. 1991.
[2] Sheldon M. Ross. A First Course in Probability. Prentice Hall. 2009.
[3] Gregory F. Lawler. Random Walk and the Heat Equation. University of Chicago Press. 2010.
[4] Peter Mörtes and Yuval Peres. Brownian Motion. Cambridge University Press. 2010.