Constructions of Brownian Motion, Reflection Principles, and the Kolmogorov-Smirnov Distribution

Mitchell Kessler

Contents

0.1 Overview 1

1 Definitions and Inequalities **1**
1.1 Notation 1
1.2 Stochastic Processes 2
1.3 Important Inequalities 2

2 Brownian Motion **5**
2.1 Definition of Brownian Motion 5
2.1.1 One-Dimensional Brownian Motion 5
2.1.2 Higher-Dimensional Brownian Motion 8
2.2 Constructions of Pre-Brownain Motion and Brownian Motion . . 9
2.2.1 Kolmogorov's Consistency Theorem 9
2.2.2 Fourier Series Expansion 10
2.2.3 Moving from Pre-Brownian Motion to Brownian Motion . 12
2.2.4 Lévy-Ciesielski's Construction 16
2.3 Brownian motion as the Limit of Distributions 20

3 Detailed Properties of Brownian Sample Paths **25**
3.1 Filtrations 25
3.2 Blumenthal's Zero-One Law 25
3.3 Zero Set 27
3.4 Nowhere Monotonicity 28
3.5 Nowhere Differentiability 28
3.6 Modulus of Continuity 29

4 The Brownian Bridge **31**
4.1 Stopping Times and the Strong Markov Property 32
4.2 Reflection Principle for Brownian Motion 34
4.3 Reflection Principles for Brownian Bridges 35

4.3.1 The Brownian Bridge as the Limit of the Empirical Process, and the Kolmogorov-Smirnov Distribution 37
4.3.2 More Bounds for Probabilities Regarding the Brownian Bridge 39

5 Appendix **42**
5.1 Kolmogorov's Inequality 42
5.2 The Central Limit Theorem and the Law of Large Numbers . . . 42

Introduction

0.1 Overview

This paper is organized into four sections, followed by an appendix.

Section 1 will begin with overview of some definitions from probability theory for clarity of presentation, along with some notation. The rest of this subsection will then be centered on stochastic processes, and will end with the Borel-Cantelli and a double-sided inequality for the standard normal distribution function.

Section 2 introduces the definition Brownian motion and some properties of Brownian motion. After this, two constructions of pre-Brownian motion will be given, followed by two methods to generate Brownian motion from pre-Brownain motion. A third construction of pre-Brownian motion, due to Lévy and Ciesielski, will be given; and by construction, this pre-Brownian motion will be sample continuous, and thus will be Brownian motion. To finish this section, a discussion of Donsker's Theorem, which shows how Brownian motion arises as the limit of a particular sequence of laws, will be given.

Section 3 presents some more detailed properties of Brownian sample paths, including its nowhere monotonicity, nowhere differentiability, and its modulus of continuity. The almost sure finiteness of hitting times will also be proven.

Section 4 is dedicated to the Brownian bridge, and giving some explicit expressions concerning its probability. Stopping times will be defined and three examples will be given, which will consequently be followed by the proof of the strong Markov property of Brownian motion. This will be used to prove the reflection principles for Brownian motion and the Brownian bridge. After this, a rather brief discussion of the appearance of the Kolmogorov-Smirnov distribution in statistics will be given, followed by an explicit expression for this distribution. To end, two other bounds for probabilities of the Brownian bridge will be proven.

The appendix consists of the statement of Kolmogorov's Inequality, as well as the statements of the Strong Law of Large Numbers and the Central Limit Theorem (both single dimensional and multivariate).

1 Definitions and Inequalities

1.1 Notation

‘

Given a collection of sets A, the σ-algebra generated by A will be denoted by $\sigma(A)$. The letter P will only be used for probability measures. The conditional probability of B with respect to A will be denoted by $P[B|A]$.

Random variables will usually be denoted by $X, Y,$ and Z. The law of a random variable X will be denoted by $\mathfrak{L}(X)$.

The set of all square-integrable functions a probability space (Ω, Σ, P) will be denoted by $L^2(\Omega, \Sigma, P)$, or just $L^2(\Omega)$ if there will be no ambiguity.

The characteristic function for the set A will be denoted by χ_A, i.e,

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

1.2 Stochastic Processes

Let (Ω, Σ, P) be a probability space.

Let T be any indexing set. T need not even be ordered, but it will be assumed to be ordered for simplicity. Furthermore, T will mainly be taken as $[0, \infty)$ or $[a, b]$, where $0 \leq a < b$; but whenever the index set is just written as T, it will be assumed that T is an arbitrary ordered set.

A collection of random variables $X = \{X_t\}_{t \in T}$, where each X_t has values in a measure space (Ω', Σ'), is called a *stochastic process*, or just a *process*. Stochastic processes could just as easily have been defined as functions

$$X : T \times \Omega \to \Omega', \ (t, \omega) \mapsto X_t(\omega) = X(t, \omega)$$

such that, for every $t \in T$, $X_t : \Omega \to \Omega', \ \omega \mapsto X_t(\omega)$ is measurable, i.e. is a random variable. For our purposes, it will be assumed that the X_t are real-valued. The process X is *sample continuous* if, for every $\omega \in \Omega$, the function $t \to X_t(\omega)$ is almost surely continuous. X is a *Gaussian process* if, for every finite $I \subset T$ and any $a_i \in \Omega, \ i \in I$, the random variable $\sum_{i \in I} a_i X_i$ is centered Gaussian. The covariance function $C : T \times T \to T$ of the process X is given by

$$C(s, t) := E[X_s X_t] - E[X_s]E[X_t].$$

In particular, if X is a Gaussian process, then $C(s, t) = E[X_s X_t]$.

The *finite dimensional distributions* of a process $X = \{X_t\}_{t \in T}$ are the distributions of the random vectors $(X_{t_1}, ..., X_{t_n})$, where $t_1 < ... < t_n \in T$ (here we're using the order assumption on T).

1.3 Important Inequalities

A few inequalities will be used throughout the text, and will be given here for reference.

Lemma 1. *(Borel-Cantelli Lemma)* Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of events in a probability space (Ω, Σ, P).

1. If $\sum_{n=1}^{\infty} P[A_n] < \infty$, then $P[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m] = 0$;
2. If the A_n are independent and $\sum_{n=1}^{\infty} P[A_n] = \infty$, then $P[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m] = 1$.

Proof. 1. Let $B_n = \bigcup_{m=n}^{\infty} A_m$. Then $B_{n+1} \subset B_n$ for all n and by measure continuity,

$$P[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m] = P[\bigcup_{n=1}^{\infty} B_n] = \lim_{n\to\infty} P[B_n].$$

Furthermore, by countable subadditivity we have

$$P[B_n] = P[\bigcup_{m=n}^{\infty} A_m] \le \sum_{m=n}^{\infty} P[A_m].$$

Since the series $\sum_{m=1}^{\infty} P[A_m]$ converges, the series $\sum_{m=n}^{\infty} P[A_m]$ converges to 0; and thus $\lim_{n\to\infty} P[B_n] = 0$.

2. Let B_n be as above. Then

$$\begin{aligned} P[\Omega - B_n] &= P[\Omega - \bigcup_{m=n}^{\infty} A_m] \\ &= P[\bigcap_{m=n}^{\infty} (\Omega - A_m)] \\ &= \prod_{m=n}^{\infty} P[\Omega - A_m] = \prod_{m=n}^{\infty} (1 - P[A_m]), \end{aligned}$$

where we have used the independence of the A_m in the third line. Now, using the fact that $1 - P[A_m] \le e^{-P[A_m]}$ (this is valid since $0 \le P[A_m] \le 1$),

$$\prod_{m=1}^{n} (1 - P[A_m]) \le \exp(-\sum_{m=1}^{n} P[A_m]).$$

By assumption, the right hand side of this expression goes to 0 as $n \to \infty$. Hence $P[\Omega - B_n] = 0$, and so $P[B_n] = 1$. This gives that $P[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m] = P[\bigcap_{n=1}^{\infty} B_n] = 1$.

$\square$

Lemma 2. For every $a > 0$,

$$(a + a^{-1})^{-1} \exp(-\frac{a^2}{2}) \le \int_a^{\infty} \exp(-\frac{x^2}{2}) dx \le a^{-1} \exp(-\frac{a^2}{2}). \tag{2}$$

Furthermore,

$$\frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-\frac{x^2}{2}} dx \le e^{-\frac{a^2}{2}}. \tag{3}$$

Proof. We have

$$\begin{aligned} \int_a^{\infty} \exp(-\frac{x^2}{2}) dx &\le \int_a^{\infty} \frac{x}{a} \exp(-\frac{x^2}{2}) dx \\ &= a^{-1} \exp(-\frac{a^2}{2}). \end{aligned}$$

For the other inequality, note that $\frac{1}{a}\exp(-\frac{a^2}{2}) = \int_a^\infty (1+\frac{1}{x^2})\exp(-\frac{x^2}{2})dx$. Thus

$$\begin{aligned}\frac{1}{a}\exp(-\frac{a^2}{2}) &= \int_a^\infty (1+\frac{1}{x^2})\exp(-\frac{x^2}{2})dx \\ &\leq (1+a^{-2})\int_a^\infty \exp(-\frac{x^2}{2})dx.\end{aligned}$$

Dividing both sides by $1+\frac{1}{a^2}$ and simplifying gives the desired inequality.

If $a > \frac{1}{\sqrt{2\pi}}$, then we can apply the inequality on the right that was just proved to get the desired inequality. If $a \leq \frac{1}{\sqrt{2\pi}}$, then we have

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_a^\infty e^{-\frac{x^2}{2}}dx &\leq \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{x^2}{2}}dx \\ &= \frac{1}{2} \leq e^{-\frac{1}{2}(\frac{1}{\sqrt{2\pi}})^2} \leq e^{-\frac{a^2}{2}}.\end{aligned}$$

□

2 Brownian Motion

2.1 Definition of Brownian Motion

2.1.1 One-Dimensional Brownian Motion

We can now state the definition of Brownian motion. The index set T will be assumed to be the set $[0, \infty)$, but everything detailed in this section is equally as valid for any closed subinterval of $[0, \infty)$.

Definition 1. *Brownian motion* is a sample continuous centered Gaussian process $B = \{B_t\}_{t \in [0,\infty)}$ such that $C(s,t) = \min\{s,t\}$ for all $s,t \in [0,\infty)$.

Remark: If B is not sample continuous, but has the other properties in this definition, then B is called *pre-Brownian motion.* Note that in order to construct Brownian motion, we will first construct pre-Brownian motion, and then will modify it in order to make it sample continuous and hence Brownian motion. Therefore, for most of this section I will be focusing on properties of pre-Brownian motion (which will themselves be valid for Brownian motion).

Pre-Brownian motion has the following useful properties:

Proposition 1. Let $B = \{B_t\}_{t \in [0,\infty)}$ be pre-Brownian motion. Then

1. $B_0 = 0$ almost surely;
2. For all $0 \leq s < t < \infty$, $B_t - B_s$ is $N(0, t-s)$;
3. For all $0 < t_1 < ... < t_n$, the increments $B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$ are independent.

In order to prove this result, we will need the following technical lemma. Here, a *Gaussian space* is defined to be any closed linear subspace of $L^2(\Omega, \Sigma, P)$ that solely consists of centered Gaussian variables.

Lemma 3. Let A be a Gaussian space and let $\{A_t : t \in T\}$ be a collection of vector subspaces of A. Then the subspaces A_t are pairwise orthogonal with respect to the covariance function $C(X,Y) = E[XY]$ on $L^2(\Omega, \Sigma, P)$ if and only if the $\sigma(A_t)$ are pairwise independent.

I will forego the proof of this lemma. It requires an application of the monotone class lemma, and the fact that, given a Gaussian vector $X = (X_1, ..., X_n)$ (i.e, any n-tuple of random variables on the same probability space (Ω, Σ, P) such that, for all $a_1, ..., a_n \in \mathbb{R}$, the random variable $\sum_{i=1}^n a_i X_i$ is centered Gaussian) has independent coordinates if and only if the covariance matrix $(C(X_i, X_j))_{i,j=1,...,n}$ is diagonal. See [3], Theorem 1.2 for the proof.

Using this, we can prove Proposition 1.

Proof. Suppose that $\{B_t\}_{t \in [0,\infty)}$ is a Gaussian process and for every $s,t \in [0,\infty]$, $E[B_s B_t] = \min\{s,t\}$. Since B_0 is $N(0,0)$, $B_0 = 0$ almost surely. This gives (1).

Fix any $s \geq 0$. Let A be the Gaussian space generated by $\{B_t\}_{t \in [0,\infty)}$, let A_s be the vector space spanned by $\{B_t\}_{t \in [0,s]}$, and let A'_s be the vector space spanned by $\{B_{s+t} - B_s\}_{t\in[0,\infty)}$. Then A_s and A'_s are both vector subspaces of the Gaussian space A. Furthermore, whenever $r \in [0, s]$,

$$\begin{aligned} E[B_r(B_{s+u} - B_s)] =& E[B_rB_{s+u}] - E[B_rB_s] \\ =& \min\{r, s+u\} - \min\{r, s\} \\ =& r - r = 0. \end{aligned} \tag{4}$$

This implies that A_s and A'_s are orthogonal. We may now apply Lemma 3 to see that the σ-algbras generated by A_s and A'_s are independent. Since B is a Gaussian process, $B_t - B_s$ is centered Gaussian; and since

$$\begin{aligned} E[(B_t - B_s)^2] &= E[B_t^2] - E[B_tB_s] - E[B_tB_s] + E[B_s^2] \\ &= t - 2\min\{s, t\} + s = t - s, \end{aligned} \tag{5}$$

it follows that $B_t - B_s$ is $N(0, t-s)$. This gives (2).

To prove (3), let $0 = t_0 < t_1 < t_2 < \cdots < t_n$ be given and set $s = t_{n-1}$ and $t = t_n$. Then $B_{t_n} - B_{t_{n-1}}$ is independent of the Gaussian vector $(B_{t_1}, ..., B_{t_{n-1}})$. Furthermore, $B(t_{n-1}) - B(t_{n-2})$ is independent of the Gaussian vector $(B_{t_1}, ..., B_{t_{n-2}})$, and so on. This implies that the variables $B_{t_i} - B_{t_{i-1}}$, $i = 1, ..., n$ are independent. □

A remark needs to be made regarding Proposition 1:

1. Any sample continuous process that satisfies properties 1 - 3 in Proposition 1 is in fact Brownian motion. To see this, note that the independent increments property implies that B is a Gaussian process. Furthermore, whenever $0 \le s < t$, we have

$$\begin{aligned} E[B_sB_t] &= E[B_s(B_t - B_s) + B_s^2] \\ &= E[B_s]E[B_t - B_s] + E[B_s^2] \\ &= 0 + s = s, \end{aligned} \tag{6}$$

so that $C(s, t) = \min\{s, t\}$. In particular, we could just as easily taken this as the definition of Brownian motion.

Using this proposition, we can compute the densities of the finite dimensional distributions of B.

Proposition 2. If $B = \{B_t : t \in [0, \infty)\}$ is a pre-Brownian motion, then for any choice of $0 = t_0 < t_1 < ... < t_n$, the law of the vector $(B_{t_1}, ..., B_{t_n})$ has density given by

$$p(x_1, ..., x_n) = \frac{1}{\sqrt{(2\pi)^n t_1(t_2 - t_1)...(t_n - t_{n-1})}} \exp\big(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\big). \tag{7}$$

where we have set $x_0 = 0$.

Proof. Assume that B is a Brownian motion. Then the variables $B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$ are independent and have $N(0, t_1), N(0, t_2 - t_1), ..., N(0, t_n - t_{n-1})$ distributions, respectively. This implies the density of the vector $(B_{t_1} - B_{t_0}, ..., B_{t_n} - B_{t_{n-1}})$ is the product of the densities of these vectors, or

$$p(u_1, ..., u_n) = \frac{1}{\sqrt{(2\pi)^n t_1 (t_2 - t_1)...(t_n - t_{n-1})}} \exp\,(-\sum_{i=1}^{n} \frac{u_i^2}{2(t_i - t_{i-1})}). \tag{8}$$

Making the change of variables $x_i = \sum_{j=0}^{i} u_j$ completes the proof.

□

Note that this proposition, together with the property that $B_0 = 0$ almost surely, characterizes the collection of finite dimensional distributions of Brownian motion. In particular, *any sample continuous stochastic process with the same finite dimensional distributions as Brownian motion is itself Brownian motion*. We will use this fact in the first construction of pre-Brownian motion

Brownian motion sample paths have the following elementary properties:

Proposition 3. Let $B = \{B_t\}_{t \in [0,\infty)}$ be a Brownian motion. Then

1. (*Symmetry*) $\{-B_t\}_{t \in [0,\infty)}$ is a Brownian motion;
2. (*Invariance under scaling*) For every $\lambda > 0$, $\{\frac{1}{\sqrt{\lambda}} B_{\lambda t}\}_{t \in [0,\infty)}$ is a Brownian motion;
3. (*Simple Markov property*) For every $s \geq 0$, $\{B_{s+t} - B_s\}_{t \in [0,\infty)}$ is a Brownian motion. Furthermore, it is independent of $\sigma(\{B_r\}_{r \in [0,s]}$; and
4. (*Time Reversal*) If $\{B_t\}_{t \in [0,T]}$ is a Brownian motion, then $\{B_T - B_{T-t}\}_{t \in [0,T]}$ is a Brownian motion.

Proof. For $\omega \in \Omega$, define $\phi_\omega(t) := B_t(\omega)$ for $t \in [0, \infty)$. It is immediate that B being a Gaussian process implies all of the listed processes are Gaussian. Thus, to show that each process is a Brownian motion, we must show the covariances are the minimum function, and that each of the processes are sample continuous. These will all be straightforward computations:

1. For every $s, t \in [0, \infty)$ we have

   $$C(s,t) = E[(-B_s)(-B_t)] = E[B_s B_t] = \min\{s, t\}.$$

   Now let $f(t) = -t$. Then f is continuous on $\mathbb{R}$; and because mapping ϕ_ω is continuous for every ω, so is the mapping $(f \circ \phi_\omega)(t) = -B_t(\omega)$.

2. For every $s, t \in [0, \infty)$,

   $$C(s,t) = E[(\frac{1}{\sqrt{\lambda}} B_{\lambda t})(\frac{1}{\sqrt{\lambda}} B_{\lambda s})] = \frac{1}{\lambda} E[B_{\lambda t} B_{\lambda s}]$$
   $$= \frac{1}{\lambda} \min\{\lambda s, \lambda t\} = \min\{s, t\}.$$

Furthermore, for every $\lambda \in [0, \infty)$, the function $f(t) = \lambda t$ is continuous, and its image on $[0, \infty)$ is $[0, \infty)$. Since ϕ_ω is continuous for every ω, it follows that $\frac{1}{\sqrt{\lambda}}(\phi_\omega \circ f)(t) = \frac{1}{\sqrt{\lambda}} B_{\lambda t}(\omega)$ is continuous.

3. For any $r, t \in [0, \infty)$, we have

$$\begin{aligned} C(r,t) &= E[(B_{s+r} - B_s)(B_{s+t} - B_s)] \\ &= E[B_{s+r}B_{s+t}] - E[B_{s+r}B_s] - E[B_sB_{s+t}] + E[(B_s)^2] \\ &= (s + \min\{r,t\}) - 2s + s = \min\{r,t\}. \end{aligned}$$

   Now note that for any fixed $s \in [0, \infty)$, the mapping $f(t) = t + s$ is continuous, and for every ω, the constant mapping $g(t) = B_s(\omega)$ is continuous. Therefore, for every ω the mapping $(\phi_\omega \circ f + g)(t) = B_{s+t}(\omega) - B_s(\omega)$ is continuous.

   To show the independence part of the statement, let A be the Gaussian space generated by the process B, and let A_s and A'_s be the vector spaces generated by $\{B_t\}_{t\in[0,s]}$ and $\{B_{s+t} - B_s\}_{t\in[0,\infty)}$, respectively. Then $\sigma(\{A_s\})$ and $\sigma(\{A'_s\})$ are independent by Proposition 1, and hence $\{B_{s+t} - B_s\}_{t\in[0,\infty)}$ is independent of $\sigma(\{B(t)\}_{t\in[0,s]})$.

4. For every $t, s \in [0, T]$, we have

$$\begin{aligned} C(s,t) &= E[(B_T - B_{T-s})(B_T - B_{T-t})] \\ &= E[(B_T)^2] - E[B_TB_{T-t}] - E[B_{T-s}B_T] + E[B_{T-s}B_{T-t}] \\ &= T - (T-t) - (T-s) + (T - \max\{s,t\}) \\ &= t + s - \max\{s,t\} = \min\{s,t\}. \end{aligned}$$

   The argument for sample continuity is the same as for the simple Markov property.

□

Further properties of the Brownian sample paths can be found in the next section of the paper.

2.1.2 Higher-Dimensional Brownian Motion

Now that we have defined Brownian motion in one dimension, we can easily define it in higher definitions. An *n-dimensional Brownian motion* is the Cartesian product of n one-dimensional Brownian motions (in the sense of the first two definitions). With this definition, we need to only study one-dimensional Brownian motions. Therefore, the constructions in the next section will solely be of one-dimensional Brownian motion.

Not only does n-dimensional Brownian motion share the same properties of one-dimensional Brownian motion, but n-dimensional Brownian motion is also independent of rotations: for any $n \times n$ rotation matrix R and any n-dimensional Browian motion and any $n \times n$ rotation matrix R, the stochastic process $B' := RB$ is also a Brownian motion.

2.2 Constructions of Pre-Brownain Motion and Brownian Motion

2.2.1 Kolmogorov's Consistency Theorem

Often times one may one to specify a stochastic process by its finite dimensional distributions. To be more specific, let T be an index set, and let $I \subset T$ be finite (say, with n elements). Suppose we are given a collection of laws P_I on $(R^n, \mathfrak{B}(\mathbb{R}^n))$ for each finite $I \subset T$. The question posed is as follows : *Does there exist a probability space* (Ω, Σ, P) *and a stochastic process* $X = \{X_t\}$ *on this space whose finite dimensional distributions are the* P_I? In general, the answer is no; but under certain consistency conditions, we can always guarantee the existence of such a probability space and such a process. The following theorem, due to Kolmogorov, gives precise conditions under which the answer to this question is affirmative.

Theorem 1. *(Kolmogorov's Consistency Theorem)* Let T be some indexing set. Suppose that for every finite subset $I := \{t_1, ..., t_n\} \subset T$, we are given a probability measure P_I on the space $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^\ltimes))$. Furthermore, suppose that these probabilities are *consistent*, in the following senses:

1. For every permutation $\pi : \{1, ..., n\} \to \{1, ..., n\}$,
$$P_{(t_1,...,t_n)} = P_{(t_{\pi(1)},...,t_{\pi(n)})};\ \text{and} \tag{9}$$
2. Whenever $I \subset J \subset T$ are finite (where J has m elements), and $B \in \mathfrak{B}(R)^n$,
$$P_I(B) = P_J(B \times \mathbb{R}^{m-n}). \tag{10}$$

Then there is a probability space (Ω, Σ, P) and a stochastic process $X := \{X_t\}_{t \in T}$ on this space whose finite dimensional distributions are precisely the $P_I,\ I \subset T$ finite.

Proof. I will give an outline of the proof. Refer to [1], Section 3 for all of the details.

Let $\mathbb{R}^T$ denote the set of all functions from T to $\mathbb{R}$. Define $\Omega := \mathbb{R}^T$.

Let A be the collection of all cylinder sets of $\mathbb{R}^T$, i.e, the collection of all sets of the form
$$C := \{B \times \mathbb{R}^{T-I} : I \subset T,\ I \text{ finite (n elements)},\ B \subset \mathfrak{B}(\mathbb{R}^n)\}. \tag{11}$$
This collection is an algebra, but it is *not* a σ-algebra. So, define $\Sigma = \sigma(A)$.

Finally, define $P : A \to [0, \infty]$ by

$$P(B \times \mathbb{R}^{T-I}) := P_I(B) \tag{12}$$

for every $I \subset T$ finite (with n elements), and every $B \in \mathfrak{B}(\mathbb{R}^n)$.

P is well-defined by the consistency conditions. It is a pre-measure on A, and by the Carathéodory-Hahn Extension Theorem, we can extend P to a probability measure (which will still be denoted by P) on Σ.

The space (Ω, Σ, P) is the probability space we will work with. On this space, define the process $X := \{X_t\}_{t \in T}$ by

$$X_t(\omega) := \omega(t) \tag{13}$$

for every $\omega \in \mathbb{R}^T$. This process satisfies the desired conditions. □

We can actually use this theorem in order to construct pre-Brownian motion. Let $T = [0, \infty)$, and for each $I = \{t_1, ..., t_n\} \subset T$ with $t_1 < t_2 < ... < t_n$, let $P_I(t) = \frac{1}{\sqrt{(2\pi)^n t_1 (t_2 - t_1) ... (t_n - t_{n-1})}} \int_{-\infty}^t \exp\left(- \sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right)$ (i.e, the P_I are just the finite dimensional distributions of pre-Brownian motion). These laws are consistent, since, for $s < t$,

$$\frac{1}{\sqrt{2\pi t}} \exp \frac{-(b-a)}{2t} = \int_{-\infty}^{\infty} \frac{\exp \frac{-(b-x)^2}{2(t-s)}}{\sqrt{2\pi(t-s)}} \cdot \frac{\exp \frac{-(x-a)^2}{2s}}{\sqrt{2\pi s}} dx.$$

By Kolmogorov's Consistency Theorem, there exists an appropriate probability space (Ω, Σ, P) and a stochastic process $X = \{X_t\}_{t \in [0, \infty)}$ on this space whose finite dimensional distributions are the P_I. In particular, X has the same finite dimensional distributions as pre-Brownian motion, and therefore is itself pre-Brownian motion.

2.2.2 Fourier Series Expansion

In this subsubsection, I will give another construction of pre-Brownian motion, using *Gaussian white noise*.

For that purpose, let $(M, \mathfrak{M})$ be a measurable space and let μ be any σ-finite measure on $(M, \mathfrak{M})$. An isometry γ from $L^2(M, \mathfrak{M}, \mu)$ onto some Gaussian space A is called a *Gaussian white noise with intensity* μ.

One question that immediately arises is whether or not there exists a Gaussian white noise with intensity μ for any given σ-finite measure μ. The answer is, in fact, in the affirmative. To be more specific, we have the following proposition.

Proposition 4. For every measure space $(M, \mathfrak{M})$ and any σ-finite measure μ on $(M, \mathfrak{M})$, there exists an appropriate probability space (Ω, Σ, P), a Gaussian space A in $L^2(\Omega, \Sigma, P)$, and a Gaussian white noise $\gamma : L^2(M, \mathfrak{M}, \mu) \to A$ with intensity μ.

Proof. First note that $L^2(M, \mathfrak{M}, \mu)$ is a Hilbert space (with usual inner product $\langle\cdot,\cdot\rangle$ given by $\langle f, g\rangle = \int_M f\bar{g}d\mu$). Let $\{f_t\}_{t\in T}$ be an orthonormal basis for $L^2(M, \mathfrak{M}, \mu)$. Then for every $f \in L^2(M, \mathfrak{M}, \mu)$, we may write

$$f = \sum_{t\in T}\langle f, f_t\rangle f_t, \tag{14}$$

where the coefficients in this expansion satisfy *Parseval's identity* with respect to the orthonormal basis $\{f_t\}_{t\in T}$:

$$\sum_{t\in T}\langle f, f_t\rangle^2 = ||f||^2 < \infty. \tag{15}$$

We now use the fact that there exists an appropriate probability space (Ω, Σ, P) and a family $X = \{X_t\}_{t\in T}$ of independent identically distributed standard normal Gaussian variables indexed by the set T. Now define A to be the Gaussian space generated by the collection X. Finally, for any $f \in L^2(M, \mathfrak{M}, \mu)$, define

$$\gamma(f) := \sum_{t\in T}\langle f, f_t\rangle X_t. \tag{16}$$

Writing out the inner product, this can be expressed as

$$\gamma(f) := \sum_{t\in T}(\int_\Omega f(x)f_t(x)d\mu(x))X_t.$$

γ is then a Gaussian white noise with intensity μ. To see this, note that the sum of the coefficients of X_t in $\gamma(f)$ is finite by Parseval's identity. Because the X_t form an orthonormal system in $L^2(\Omega, \Sigma, P)$, the sum on the right-hand side above actually converges in $L^2(\Omega, \Sigma, P)$. By definition, γ takes values in the Gaussian space A. Fixing any $t_1 \in T$ and using the orthonormality of the system $\{f_t\}_{t\in T}$, it follows that

$$\begin{aligned}\gamma(f_{t_1}) &= \sum_{t\in T}\langle f_t, f_{t_1}\rangle X_t \\ &= \langle f_{t_1}, f_{t_1}\rangle X_{t_1} \\ &= X_{t_1}.\end{aligned}$$

This implies that γ is an isometry from $L^2(\Omega)$ to $A = \text{span}(X_t : t \in T)$. Thus γ is Gaussian white-noise with intensity μ. □

Now let γ be any Gaussian white noise with intensity μ, defined on $L^2(M, \mathfrak{M}, \mu)$ and with values in $L^2(\Omega, \Sigma, P)$. Then:

1. For every $f \in L^2(M, \mathfrak{M}, \mu)$, $\gamma(f)$ is a centered Gaussian variable, with variance

$$E[\gamma(f)^2] = ||\gamma(f)||^2_{L^2(\Omega,\Sigma,P)} = ||f||^2_{L^2(M,\mathfrak{M},\mu)} = \int f^2 d\mu; \text{ and} \tag{17}$$

2. for every $f, g \in L^2(M, \mathfrak{M}, \mu)$, we have

$$\begin{aligned} E[\gamma(f)\gamma(g)] &= \langle \gamma(f), \gamma(g) \rangle_{L^2(\Omega,\Sigma,P)} \\ &= \langle f, g \rangle_{L^2(M,\mathfrak{M},\mu)} = \int fg d\mu. \end{aligned} \tag{18}$$

With the existence of Gaussian white noise in mind, we can easily construct pre-Brownian motion. Consider the measure space $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mu)$, where μ is Lebesgue measure. From the above remarks, there exists an appropriate probability space (Ω, Σ, P) and a Gaussian white noise γ with intensity μ, defined on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mu)$ and with values in (Ω, Σ, P). For every $t \in [0, \infty)$, define

$$X_t := \gamma(\chi_{[0,t]}). \tag{19}$$

Because the $\gamma(\chi_{[0,t]})$ belong to some common Gaussian space, it follows immediately that $X = \{X_t\}_{t\in[0,\infty)}$ is a Gaussian process. Furthermore, for any $s, t \in [0, \infty)$, we have

$$E[X_t X_s] = E[\gamma(\chi_{[0,t]})\gamma(\chi_{[0,s]})] = \int_0^\infty \chi_{[0,t]}(x)\chi_{[0,s]}(x) d\mu(x) = \min\{s, t\}.$$

Hence X is pre-Brownian motion.

2.2.3 Moving from Pre-Brownian Motion to Brownian Motion

Here, I will give two separate methods from moving from pre-Brownian motion to Brownian motion. The first method heavily relies on the Gaussian properties of pre-Brownian motion, while the second method generates a modification of pre-Brownian motion that is actually Hölder continuous.

Method 1: Let $X = \{X_t\}_{t\in[0,\infty)}$ be any pre-Brownian motion. It suffices to construct a stochastic process with the same finite-dimensional distributions as X.

To do this, for every $n \geq 1$, define

$$Y_k := X_{\frac{k+1}{2^n}} - X_{\frac{k}{2^n}}, \tag{20}$$

where $k = 0, 1, ..., 2^n - 1$. Since X is pre-Brownian motion, Y_k is $N(0, \frac{1}{2^{-n}})$. By Lemma 2, we have

$$P[\sup_k |Y_k| \geq \frac{1}{n^2}] \leq 2^n P[|Y_1| \geq \frac{1}{n^2}] \leq 2^{n+1} \exp(-\frac{2^{n-1}}{n^4}). \tag{21}$$

The right hand side of this expression is the general term of a convergent series. By the Borel-Cantelli Lemma,

$$P[\bigcap_{m\geq 1} \bigcup_{k\geq m} \{\sup_k |Y_k| \geq \frac{1}{n^2}\}] = 0. \tag{22}$$

For any $t \in [0,1]$, consider its dyadic expansion $t := \sum_{i=1}^{\infty} \frac{t_i}{2^i}$, where $t_i = 0, 1$ for every i. For every $n \geq 1$, define

$$t_n := \sum_{i=1}^{n} \frac{t_i}{2^i}. \tag{23}$$

This definition implies that $X_{t_n} - X_{t_{n-1}} \in \{0, Y_1, ..., Y_{2^n-1}\}$ for every $n \geq 1$. Noticing that $X_{t_n} = \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})$, and using the fact that $P[\bigcap_{m\geq 1} \bigcup_{k\geq m} \{\sup_k |Y_k| \geq \frac{1}{n^2}\}] = 0$, we see that there exists an $N > 0$ such that $n \geq N$ implies

$$P[\{|X_{t_n} - X_{t_{n-1}}| \leq \frac{1}{n^2}\}] = 1. \tag{24}$$

Therefore, the sequence $\{X_{t_n}\}_{n=1}^{\infty}$ converges almost surely to some limit Z_t on $[0,1]$. We will show that Z is the desired sample continuous process.

By construction, $Z_t = X_t$ for any dyadic $t \in [0,1]$. Now, choose any $t, s \in [0,1]$ with $|t - s| \leq 2^{-n}$. If in fact $t = t_n = \frac{k}{2^n}$ and $s = s_n = \frac{m}{2^n}$, then $|k - m| = 0, 1$. In particular, either $|X_{t_n} - X_{s_n}| = 0$; or $|X_{t_n} - X_{s_n}| = |Z_k|$ for some k; and in this case we have $|X_{t_n} - X_{s_n}| \leq \frac{1}{n^2}$ for large enough n. By definition of Z_t, it follows that for large enough n,

$$\begin{aligned} |Z_t - Z_s| &\leq |Z_t - X_{t_n}| + |X_{t_n} - X_{s_n}| + |X_{s_n} - Z_s| \\ &\leq \sum_{m=n}^{\infty} \frac{1}{m^2} + \frac{1}{n^2} + \sum_{m=n}^{\infty} \frac{1}{m^2} \\ &\leq \frac{C}{n} \end{aligned} \tag{25}$$

for some constant C. Thus $t \mapsto Z_t(\omega)$ is continuous for all $\omega \in \{|X_{t_n} - X_{t_{n-1}}| \leq \frac{1}{n^2}\}$. Since this set has probability one, it follows that Z is sample continuous. On the complement of this set, set $Z_t = 0$. Using the sample continuity and the fact that Z agrees with X on a dense subset of $[0,1]$, we can conclude that Z has the same finite-dimensional distributions as Brownian motion. In particular, Z is itself Brownian motion.

Method 2: For this method, we will first need to give two definitions.

Definition 2. Let $X = \{X_t\}_{t\in T}$ and $Y = \{Y_t\}_{t\in T}$ be stochastic processes defined on the same index set T and with values in a mutual measure space $(M, \mathfrak{M})$. Then Y is a *modification* of X if for every $t \in T$, $P(\{Y_t = X_t\}) = 1$; and Y is *indistinguishable* from X if $P(\{X_t = Y_t, \forall t \in T\}) = 1$, or, equivalently, if there is a subset $N \subset \Omega$ of measure zero such that $X_t = Y_t$ whenever $\omega \in \Omega - N$.

The goal of this section is to show that there exists a suitable modification of pre-Brownian motion that is sample continuous. In fact, we will show something much stronger: namely, it will be shown that every pre-Brownian motion can be modified to be Hölder continuous. This is the contents of the following theorem.

Theorem 2. Let $X = \{X_t\}_{t\in[0,\infty)}$ be a pre-Brownian motion. There exists a modification of X that is sample continuous, and is in fact locally Hölder continuous with exponent $\frac{1}{2} - \eta$ whenever $\eta \in (0, \frac{1}{2})$.

Proof. We need two lemmas in order to prove this result.

Lemma 4. Let (E, d) be any metric space, let $D = \{\frac{k}{2^n} : n \in \mathbb{N}, 1 \leq k \leq 2^n - 1\}$, and let $f : D \to E$ be a mapping. Assume that there exists an $\alpha > 0$ and some finite constant C such that for every $n \in \mathbb{N}$ and for every $k \in \mathbb{N}$ with $1 \leq k \leq 2^n - 1$,

$$d(f((i-1)2^{-n}), f(i2^{-n})) \leq C2^{-n\alpha}. \tag{26}$$

Then whenever $s, t \in D$,

$$d(f(s), f(t)) \leq \frac{C}{1 - 2^{-\alpha}}|t - s|^{\alpha}. \tag{27}$$

Proof. Fix any $s, t \in D$ with $s < t$. Let n be the smallest positive integer such that $2^{-n} \leq t - s$, and let m be the smallest nonnegative integer such that $m2^{-n} \geq s$. Then we can write

$$s = m2^{-n} - e_0 2^{-n-1} - ... - e_k 2^{-n-k}; \text{ and} \tag{28}$$

$$t = m2^{-n} + e'_0 2^{-n-1} + ... + e'_p 2^{-n-p}, \tag{29}$$

for some nonnegative integers k, p and constants e'_i, e_j that are either 0 or 1 for every $i = 0, ..., p$ and $j = 0, ..., k$. Furthermore, define

$$s_j := m2^{-n} - e_0 2^{-n-1} - ... - e_j 2^{-n-j}; \text{ and} \tag{30}$$

$$t_i := m2^{-n} + e'_0 2^{-n-1} + ... + e'_i 2^{-n-i}. \tag{31}$$

Since $t_p = t$ and $s_k = s$, we can apply the assumptions of the Lemma to each of the pairs $(s_0, t_0), (s_{j-1}, s_j)$, and (t_{i-1}, t_i), to get

$$\begin{aligned}
d(f(s), f(t)) &= d(f(s_k), f(t_p)) \\
&\leq d(f(s_0, t_0)) + \sum_{j=1}^{k} d(f(s_j), f(s_{j-1})) + \sum_{i=1}^{p} d(f(t_i), f(t_{i-1})) \\
&\leq C2^{-n\alpha}(1 + \sum_{j=1}^{k} 2^{-j\alpha} + \sum_{i=1}^{p} 2^{-i\alpha}) \\
&\leq C2^{-n\alpha}(1 + 2\sum_{i=1}^{\infty} 2^{-i\alpha}) \\
&\leq 2C(1 - 2^{-\alpha})^{-1}2^{-n\alpha} \\
&\leq 2C(1 - 2^{-\alpha})^{-1}(t - s)^{\alpha}.
\end{aligned} \tag{32}$$

The second to last inequality follows from the fact that $(1 + 2\sum_{i=1}^{\infty} 2^{-i\alpha}) < 2(\sum_{i=0}^{\infty} 2^{-i\alpha}) = \frac{2}{1-2^{-\alpha}}$; and the last line follows from $2^{-n} \leq t - s$. $\square$

Lemma 5. *(Kolmogorov's Lemma)* Let $X = \{X_t\}_{t \in T}$ be a stochastic process indexed by some real bounded interval T, where all the X_t take values on some complete metric space (E, d). Assume that there are three real numbers $q, \epsilon, C > 0$ such that whenever $s, t \in T$,

$$E[d(X_s, X_t)^q] \leq C|t - s|^{1+\epsilon}. \tag{33}$$

Then there exists a modification Y of X whose paths are Hölder continuous with exponent $\alpha \in (0, \frac{\epsilon}{q})$, i.e, for every $\omega \in \Omega$ and every $\alpha \in (0, \frac{\epsilon}{q})$, there exists a constant $C_\alpha(\omega) < \infty$ such that for every $s, t \in T$,

$$d(Y_s(\omega), Y_t(\omega)) \leq C_\beta(\omega)|t - s|^\beta. \tag{34}$$

In particular, Y has continuous sample paths.

Proof. For simplicity, take $T = [0, 1]$; and fix any $\alpha \in (0, \frac{\epsilon}{q})$. Our assumptions and Chebyshev's Inequality imply, for any $a > 0$ and $s, t \in T$, that

$$P[d(X_s, X_t) \geq a] = P[d(X_s, X_t)^q \geq a^q] \leq \frac{E[d(X_s, X_t)^q]}{a^q} \leq Ca^{-q}|t - s|^{1+\epsilon}.$$

Applying this inequality to $s = (i - 1)2^{-n}$, $t = i2^{-n}$, and $a = 2^{-n\alpha}$ for $i = 1, ..., 2^n$, we get

$$P[d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geq a] \leq C2^{n(q\alpha-(1+\epsilon))}.$$

Summing over all i, it follows that

$$P[\bigcup_{i=1}^{2^n}\{d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geq a\}] \leq 2^n(C2^{n(q\alpha-(1+\epsilon))}) = C2^{-n(\epsilon-q\alpha)}.$$

By assumption, $\epsilon - q\alpha > 0$, from which is follows that the right-hand side of the above expression is the general term of an absolutely convergent series. Summing over all n, we then have

$$\sum_{n=1}^{\infty} P[\bigcup_{i=1}^{2^n}\{d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geq a\}] \leq C\sum_{n=1}^{\infty} 2^{-n(\epsilon-q\alpha)} < \infty.$$

By the Borel-Cantelli lemma, there almost surely exists an $N_0 = N_0(\omega)$ such that whenever $N \geq N_0$ and whenever $i = 1, 2, ..., 2^n$,

$$d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \leq 2^{-n\alpha}.$$

In particular, the constant

$$K_\alpha(\omega) := \sup_{n \geq 1}[\sup_{1 \leq i \leq 2^n} \frac{d(X_{(i-1)2^{-n}}, X_{i2^{-n}})}{2^{-n\alpha}}] \tag{35}$$

is almost surely finite (i.e, $P[\{K_\alpha < \infty\}] = 1$), since whenever $N \geq N_0$ the supremum inside parentheses is bounded above by 1; and since there are only finitely many terms less than N_0. Applying Lemma 4 on the event $\{K_\alpha < \infty\}$, we see that for every $s, t \in D$,

$$d(X_s, X_t) \leq C_\alpha(\omega)|t-s|^\alpha, \tag{36}$$

where $C_\alpha(\omega) = 2(1-2^{-\alpha})K_\alpha(\omega)$. Thus, on the event $\{K_\alpha(\omega) < \infty\}$, the sample path $t \to X_t(\omega)$ is Hölder continuous on D, and therefore is uniformly continuous on D. Because (E, d) is a complete metric space, there is a unique continuous extension of this mapping to the interval $T = [0, 1]$, which also is Hölder continuous with exponent α.

Now fix some $x_0 \in E$. Define the process $Y := \{Y_t\}_{t\in[0,1]}$ by

$$Y_t(\omega) := \begin{cases} \lim_{s\to t,\ s\in D} X_s(\omega), & \text{if } K_\alpha(\omega) < \infty; \\ x_0, & \text{if } K_\alpha(\omega) = \infty. \end{cases} \tag{37}$$

It is evident that Y_t is a random variable for each $t \in [0, 1]$. Furthermore, the comments above imply that the sample paths of Y are Hölder continuous with exponent α. Moreover, whenever $t \in [0, 1]$,

$$X_s \to X_t \tag{38}$$

in probability as $s \to t$ for $s \in D$. Since Y_t is defined as the almost sure limit of X_s as $s \to t$, $s \in T$, this implies that $X_t = Y_t$ almost surely. In particular, Y is indeed a modification of X. □

Having established these two lemmas, the theorem can be proved quickly. If $s < t$, then because the random variable $X_t - X_s$ is $N(0, t-s)$, there is some random variable Y that is $N(0, 1)$ such that $X_t - X_s = \sqrt{t-s}Y$. Therefore, whenever $\alpha > 0$,

$$E[|X_t - X_s|^\alpha] = (t-s)^{\frac{\alpha}{2}} E[|Y|^\alpha] < \infty.$$

Whenever $\alpha > 2$, we may apply Lemma 5 with $\epsilon = \frac{\alpha}{2}$ to generate a modification of X that is sample continuous and, in fact, has sample paths which are locally Hölder continuous with exponent β for any $\beta < \frac{\alpha-2}{2\alpha}$. Letting α tend to ∞ allows us to take β arbitrarily close to $\frac{1}{2}$. □

By construction, the modification Y of the stochastic process X in the proof is unique up to indistinguishability.

2.2.4 Lévy-Ciesielski's Construction

This construction is due to Lévy, but was simplified by Ciesielski. I will give Ciesielski's simplified version of the construction. Ultimately, this construction is akin to the previous construction of pre-Brownian motion (as both use Fourier series in order to construct the pre-Brownain motion), but the process constructed will actually be sample continuous by definition (this we still will have to show).

The goal of this section is to first construct Brownian motion with index set $[0,1]$, and then extend this construction to generate Brownian motion with index set $[0,\infty)$.

Let (Ω,Σ,P) be the probability space constructed in Kolmogorov's Theorem (except, assume that the original index set is $T=[0,1]$ instead of $T=[0,\infty)$), so that the P_I are just the finite dimensional distributions of pre-Brownian motion.

Let $n\geq 1$ be given, and let $1\leq k<2^n$ be odd. The *Haar functions*, defined on $[0,1]$, are given by

$$f_{k2^{-n}}(t)=\begin{cases}2^{\frac{n-1}{2}}, & \text{if } (k-1)2^{-n}\leq t<k2^{-n};\\ -2^{\frac{n-1}{2}}, & \text{if } k2^{-n}\leq t<(k+1)2^{-n};\\ 0, & \text{otherwise,}\end{cases} \tag{39}$$

together with the function $f_0\equiv 1$. This collection is a complete orthonormal system in $L^2([0,1])$, and hence is a basis. To see this, note that

$$\int_0^1 f_{i2^{-n}}(t)f_{j2^{-m}}(t)dt=\int_{(i-1)2^{-n}}^{(i+1)2^{-n}} 2^{n-1}dt=1. \tag{40}$$

If instead $i2^{-n}\neq j2^{-m}$, then either $n\neq m$ or $n=m$ but $i\neq q$. In the second case, the intervals on which the Haar functions are nonzero do not overlap, and therefore the integral evaluates to 0. So now assume $n\neq m$; without loss of generality we may assume that $n<m$. Then either $[\frac{j-1}{2^m},\frac{j+1}{2^m})\subset[\frac{i-1}{2^n},\frac{i+1}{2^n})$, or these two intervals are disjoint. If they're disjoint, we're done. Otherwise, the first interval must be contained in one of the halves of the second interval, since it's length is less than $\frac{1}{2}$ of the length of the second interval. In particular, $f_{j2^{-m}}$ is constant on the support of $f_{i2^{-n}}$, which gives

$$\int_0^1 f_{j2^{-m}}f_{i2^{-n}}=C\int_0^1 f_{i2^{-n}}=0.$$

It is also complete: if $g\in L^2([0,1])$ is perpendicular to every Haar function, then the integral

$$\int_{(k)2^{-n}}^{(k+1)2^{-n}} g(t)dt$$

is independent of off $0\leq k<2^n$. To prove this, for any integer $n\in\mathbb{N}$ and $k=0,1,...,2^n-1$, define

$$I_k:=\int_{k2^{-n}}^{(k+1)2^{-n}} g(x)dx.$$

Our assumptions on the $f_{k2^{-n}}$ imply that $I_k=I_0$ for every k, and therefore the integral above is independent of k. In particular, we have

$$\int_{i2^{-n}}^{j2^{-n}} g(t)dt=2^{-n}(j-i)\int_0^1 g(t)dt=0$$

for every $0 \le i < j \le 2^n$ and any $n \ge 1$. Thus

$$\int_a^b g(x)dx = 0.$$

holds for all dyadic rationals $a, b \in [0,1]$. This is enough to show that the Haar system is complete.

We can now begin out construction of Brownian motion with index set $[0,1]$. Let $\{g_{k2^{-n}} : k = 0 \text{ or odd } k < 2^n, 1 \le n\}$ be any collection of independent identically distributed standard Gaussian variables (they can be on any space); and for each $t \in [0,1]$ define

$$B_t := g_0 \int_0^t f_0(x)dx + \sum_{n=1}^{\infty} \sum_{k \text{ odd}<2^{-n}} g_{k2^{-n}} \int_0^t f_{k2^{-n}}(x)dx. \tag{41}$$

If the right-hand side of the B_t converge, then it immediately follows that $B = \{B_t\}_{t\in[0,1]}$ is a Gaussian process by assumption on the $g_{k2^{-n}}$.

Therefore, the first goal is to show that this sum converges on $[0,1]$, and in fact converges uniformly on $[0,1]$ to a continuous path.

To do this, consider the Schauder functions $S_{k2^{-n}}(t) = \int_0^t f_{k2^{-n}}(x)dx$. These are just little tents of height $2^{-\frac{n+1}{2}}$. With this in mind, it follows that

$$\begin{aligned} e_n &= \Big|\Big| \sum_{k \text{ odd}<2^n} g_{k2^{-n}} \int_0^t f_{k2^{-n}}(x)dx \Big|\Big|_\infty \\ &= \max_{t\in[0,1]} \Big| \sum_{k \text{ odd}<2^n} g_{k2^{-n}} \int_0^t f_{k2^{-n}}(x)dx \Big| \cdot \\ &= 2^{-\frac{n-1}{2}} \max_{k \text{ odd}<2^n} |g_{k2^{-n}}| \end{aligned} \tag{42}$$

Therefore, for any constant $C > 0$, we have

$$\begin{aligned} P[e_n > C(2^{-n}\log(2^{-n}))^{\frac{1}{2}}] &= P[\max_{k \text{ odd } <2^{-n}} g_{k2^{-n}} > C\sqrt{2n\log(2)}\ ] \\ &\le 2 \cdot 2^{n-1} \int_{C(\sqrt{2n\log(2)})}^1 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \\ &< n^{-\frac{1}{2}} 2^n K e^{-C^2 n \log(2)} \\ &= n^{-\frac{1}{2}} 2^{n(1-C^2)} K, \end{aligned} \tag{43}$$

where $K > 0$ is constant, and where we have used Lemma 2 in the third inequality. Whenever $C > 1$, the series $\sum_{n=1}^{\infty} n^{-\frac{1}{2}} 2^{n(1-C^2)}$ converges absolutely; so by the Borel-Cantelli Lemma,

$$P[\limsup_{n\to\infty} e_n \le C(2^{-n}\log(2^{-n}))^{\frac{1}{2}}] = 1. \tag{44}$$

This implies that the series converges absolutely on $[0,1]$ and that the process $B = \{B_t : t \in [0,1]\}$ is actually sample continuous.

In order to finish the proof, we need to show that $E[B(t)B(s)] = \min\{s,t\}$ whenever $0 \le s,t \le 1$. For every such s,t, we have

$$\begin{aligned} E[B(t)B(s)] &= E[(g_0 \int_0^t f_0(x)dx + \sum_{n=1}^{\infty} \sum_{k \text{ odd}<2^{-n}} g_{k2^{-n}} \int_0^t f_{k2^{-n}}(x)dx) \\ &\quad \times (g_0 \int_0^s f_0(x)dx + \sum_{n=1}^{\infty} \sum_{k \text{ odd}<2^{-n}} g_{k2^{-n}} \int_0^s f_{k2^{-n}}(x)dx)] \\ &= \int_0^t f_0(x)dx \int_0^s f_0(x)dx + \sum_{n=1}^{\infty} (\int_0^t f_{k2^{-n}}(x)dx)(\int_0^s f_{k2^{-n}}(x)dx) \\ &= \int_0^1 \chi_{[0,t]}(x)\chi_{[0,s]}(x)dx = \min\{s,t\}. \end{aligned} \tag{45}$$

Here we used the independence of the $g_{k2^{-n}}$ to get the second inequality; and to get the third equality, we have used Parseval's Inequality with respect to the the Haar orthonormal system applied to the characteristic functions $\chi_{[0,t]}$ and $\chi_{[0,s]}$. B is therefore a Brownian motion on $[0,1]$.

We can use this to construct a Brownian motion on $[0,\infty)$. Use this construction of Brownian motion to construct some sequence $\{B^n\}_{n=1}^{\infty}$ of independent copies of a Brownian motion. Define the stochastic process $B := \{B_t\}_{t\in[0,\infty)}$ by

$$B_t := \begin{cases} B_t^1, & \text{if } 0 \le t < 1; \\ B_1^1 + B_{t-1}^2, & \text{if } 1 \le t < 2; \\ \cdots\cdots\cdots \\ B_1^1 + \cdots + B_1^n + B_{t-n}^{n+1}, & \text{if } n \le t < n+1 \\ \cdots\cdots\cdots \end{cases} \tag{46}$$

This is sample continuous, since each of the B^n are sample continuous and $\lim_{t\to n} B_t = B_1^1 + ... + B_1^n$ regardless of the direction that t approaches n. It is indeed a Gaussian process, since the B^n are Gaussian processes and are all mutually independent. It remains to show that this has the desired covariance. To do this, let $0 \le s < t$ be given. We have two cases:

1. $s,\ t \in [j, j+1)$ for some $j \ge 1 \in \mathbb{Z}$. Then $B_t - B_s = B_{t-j}^{j+1} - B_{s-j}^{j+1}$, which implies

$$\begin{aligned} E[B_tB_s] &= E[(B_{t-j}^{j+1} + B_1^j + ... + B_1^1)(B_{s-j}^{j+1} + B_1^j + ... + B_1^1)] \\ &= E[B_{t-j}^{j+1}B_{s-j}^{j+1}] + \sum_{k=1}^{j} E[(B_1^k)^2] \\ &= (s-j) + j = s. \end{aligned}$$

2. $s \in [j, j+1)$ and $t \in [i, i+1)$, where $1 \le j < i$; $i,\ j \in \mathbb{Z}$. Then $B_t - B_s = B_{t-i}^{i+1} + B_1^i + ... + B_1^{j+1} - B_{s-j}^{j+1}$. Furthermore, using the independence of the

B^n, we have

$$\begin{aligned}E[B_tB_s] &= E[(B_{t-j}^{j+1} + B_1^j + ... + B_1^1)(B_{s-i}^{i+1} + B_1^i + ... + B_1^1]\\&= E[B_{s-i}^{i+1}] + \sum_{k=1}^{i} E[(B_1^k)^2]\\&= (s-i) + i = s\end{aligned}$$

B then has the desired covariance, and hence is Brownian motion.

This type of extension makes it very easy to move from a Brownian motion with index set $[0,1]$ to a Brownian motion with index set $[0,\infty)$, and allows us to solely focus on Brownian motions with index set $[0,1]$. Indeed, we also could have used this after the first construction of Brownian motion with index set $[0,1]$ in order to extend it to a Brownian motion with index set $[0,\infty)$.

2.3 Brownian motion as the Limit of Distributions

In this section, I will discuss the Donsker Theorem, which is a classical central limit theorem on $C[0,\infty)$.

Define the metric $d : C[0,\infty) \to [0,\infty)$ by

$$d(f,g) := \sup_{n=1} \frac{1}{2^n} \frac{d_n(f,g)}{1+d_n(f,g)}, \tag{47}$$

where $d_n(f,g) = \sup_{t\in[0,n]}|f(t) - g(t)|$. This function is indeed a metric, and actually metricizes uniform convergence on compact sets in $C[0,\infty)$.

Consider any sequence $(X_n)_{n=1}^{\infty}$ of independent identically-distributed random variables, each with mean 0 and finite variance σ^2. For every $n \in \mathbb{Z}_{\geq 1}$ and every $t \in [0,\infty)$, let $\lfloor nt \rfloor$ be the smallest integer less than or equal to nt, and define

$$B_t^n := \frac{1}{\sigma\sqrt{n}}(\sum_{i\leq\lfloor nt\rfloor} X_i) + (nt - \lfloor nt\rfloor)\frac{X_{1+\lfloor nt\rfloor}}{\sigma\sqrt{n}}. \tag{48}$$

Set $B^n := \{B_t^n\}_{t\in[0,\infty)}$. Simply put, B^n is a random walk at integer values of nt, and is just linear interpolated for all other values. The goal of this section is to prove the following theorem:

Theorem 3. *(Donsker's Theorem)* The stochastic process $B^n = \{B_t^n\}_{t\in[0,\infty)}$ converges in distribution to Brownian motion B_t on the metric space $(C([\infty)), d)$ as $n \to \infty$. In particular, for every $t \in [0,\infty)$,

$$\mathfrak{L}(B_t^n) \to \mathfrak{L}(B_t).$$

The proof of this theorem requires some work. First of all, probabilities on $C[0,\infty)$ are completely determined by the finite dimensional distributions. In particular, the cylindrical σ-algebra on $C[0,\infty)$ (namely, the σ-algebra generated by the events $\{f \in C[0,\infty) : (f(t_1), ..., f(t_n)) \in B,\ B \in \mathfrak{B}(\mathbb{R}^n)\}$ for some fixed B and fixed $t_1 < ... < t_n$ in $[0, \infty)$) coincides as the Borel σ-algebra on $C[0, \infty)$. This is a rather simple consequence of $C[0, \infty)$ being a separable metric space. This implies that, if the finite dimensional distributions of the sequence of stochastic processes B^n converge to the finite dimensional distributions to some stochastic process B, then B is the unique possible distributional limit of this sequence.

Now notice that because the second term in the expression for B_t^n is of the order $n^{-\frac{1}{2}}$, we can simplify our argument by treating tn as an integer and writing $B_t^n = \frac{1}{\sigma\sqrt{n}} \sum_{i \leq nt} X_i$. Writing

$$\frac{1}{\sigma\sqrt{n}} \sum_{i \leq nt} X_i = \frac{\sqrt{t}}{\sigma\sqrt{nt}} \sum_{i \leq nt} X_i,$$

we may apply the Central Limit Theorem to see that $B^n(t)$ converges in distribution to $N(0, t)$. Furthermore, whenever $s < t$, we can express B_t^n as the sum

$$B_t^n = B_s^n + \frac{1}{\sigma\sqrt{n}} \sum_{ns < i \leq nt} X_i.$$

Because B_s^n and $B_t^n - B_t^n$ are independent, it follows that

$$\begin{aligned} E[B_s^n B_t^n] &= E[B_s^n(B_t^n - B_s^n)] + E[(B_s^n)^2] \\ &= E[(B_s^n)^2], \end{aligned}$$

which converges to s as $n \to \infty$. It follows that the finite dimensional distributions of B^n converge to the finite dimensional distributions of a Brownian motion $B = \{B_t\}_{t \in [0,\infty)}$, and therefore, by the previous discussion, identifies the Brownian motion $B = \{B_t\}$ as the unique possible limit of the sequence $\{B^n\}_{n=1}^{\infty}$.

It therefore remains to show that the sequence actually converges. In order to do this, we will need the Selection Theorem, which concerns uniform tight sequences of laws on probability spaces.

To be specific, let $(P_n)_{n=1}^{\infty}$ be any sequence of laws on a metric space (E, d). We say that $(P_n)_{n=1}^{\infty}$ is *uniformly tight* if for every $\epsilon > 0$, there exists some compact set $K \subset E$ such that for every n, $P_n[K] \geq 1 - \epsilon$. We have the following theorem:

Theorem 4. *(Selection Theorem)* For any uniformly tight sequence of laws $(P_n)_{n=1}^{\infty}$ on the metric space (E, d), there exists a subsequence $(n_k)_{k=1}^{\infty}$ such that $(P_{n_k})_{k=1}^{\infty}$ converges weakly to some probability law P.

From this, it follows that if we can show that the sequence of laws $\{\mathfrak{L}(B_t^n)\}_{n=1}^{\infty}$ is uniformly tight for each t, then there will exist a subsequence $\{\mathfrak{L}(B_t^{n_k})\}_{k=1}^{\infty}$ that converges weakly to some probability law P. But the unique possible limit of this sequence will necessarily be $\mathfrak{L}(B_t)$; and the existence of one convergent subsequence, together with the fact that the finite dimensional distributions converge and identify a unique possible limit, will imply that $\{\mathfrak{L}(B_t^n)\}_{n=1}^{\infty} \to \mathfrak{L}(B_t)$ weakly in $(C[0,\infty), d)$, which will complete the proof of the Theorem.

The rest of this section will therefore be devoted to finding a convenient characterization of uniform tightness.

For any given $\delta, T > 0$ and any given function $f \in C([0,T])$, define

$$m^T(f,\delta) := \sup\{|f(b) - f(a)| : |b-a| \leq \delta,\ a, b \in [0,T]\}. \tag{49}$$

m^T is just the modulus of continuity for the function f. We have the following version of the Arzelá-Ascoli Theorem using the aforementioned notation:

Theorem 5. *(Arzelá-Ascoli Theorem)* The set K is compact in $(C([0,\infty)), d)$ if and only if for every $T > 0$,

$$\sup_{f\in K} |f(0)| < \infty \text{ and } \lim_{\delta\to 0} \sup_{f\in K} m^T(f,\delta) = 0. \tag{50}$$

We can use this to prove the following result to give an equivalent statement of uniform tightness.

Proposition 5. A sequence of laws $(P_n)_{n=1}^{\infty}$ on $(C((0,\infty]), d)$ is uniformly tight if and only if

1. $$\lim_{a\to\infty} \sup_{n\geq 1} P_n[\{f \in C((0,\infty]) : |f(0)| > a\}] = 0; \text{ and} \tag{51}$$

2. for every $T, \epsilon > 0$, we have

$$\lim_{\delta\to 0} \sup_{n\geq 1} P_n[\{f \in C((0,\infty]) : m^T(f,\delta) > \epsilon\}] = 0. \tag{52}$$

For proofs of these statements, see [1], Section 22.

The second condition in this statement is usually difficult to show, so we instead will use the following *asymptotic equicontinuity condition*:

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P_n[\{f \in C((0,\infty]) : m^T(f,\delta) > \epsilon\}] = 0. \tag{53}$$

This statement implies the second statement in Proposition 5. To see this, suppose this holds for every $T > 0$ and every $\epsilon > 0$. Then for every $a > 0$ there exists a $\delta_0 > 0$ and a $N \in \mathbb{N}$ such that whenever $n \geq N$,

$$P_n[\{f \in C((0,\infty]) : m^T(f,\delta) > \epsilon\}] < a.$$

Since every $f \in C((0,\infty])$ is continuous by definition, the modulus of continuity $m^T(f,\delta)$ goes to 0 as $\delta \to 0$. Furthermore, m^T is a decreasing function of δ. Therefore, for every $n < N$, there exists a $\delta_n > 0$ such that

$$P_n[\{f \in C((0,\infty]) : m^T(f,\delta_n) > \epsilon\}] < a;$$

and whenever $n \geq 1$ and $\delta < \min\{\delta_0, \delta_1, ..., \delta_n\}$, we have

$$P_n[\{f \in C((0,\infty]) : m^T(f,\delta) > \epsilon\}] < a.$$

In particular, $\lim_{\delta\to 0}\sup_{n\geq 1} P_n[\{f \in C((0,\infty]) : m^T(f,\delta) > \epsilon\}] = 0$ holds for every $\epsilon > 0$ and every $T > 0$.

We can now complete the proof of the uniform tightness of the sequence $\{\mathbb{L}(B_t^n)\}_{n=1}^{\infty}$ using Proposition 5. By definition, $B_0^n = 0$ almost surely, so it suffices to check that the asymptotic equicontinuity condition holds for this sequence.

Note that whenever $s < t$,

$$\begin{aligned} m^T(B^n,\delta) &= \sup_{t,s\in[0,T];\ |t-s|\leq\delta} |\frac{1}{\sigma\sqrt{n}} \sum_{ns\leq t\leq nt} X_i| \\ &\leq \max_{0\leq k\leq nT;\ 0<j\leq n\delta} |\frac{1}{\sigma\sqrt{n}} \sum_{k<i\leq k+j} X_i|. \end{aligned} \tag{54}$$

To maximize this expression over all $0 \leq k \leq nT$, we can set $m = \frac{T}{\delta}$ and can maximize over all indices $k := ln\delta$, where $0 \leq l \leq m-1$. Since we also need to maximize over $0 < j \leq n\delta$, and since both of the maximums being considered are taken over increments of size $n\delta$, it follows that

$$\max_{0\leq k\leq nT;\ 0<j\leq n\delta} |\frac{1}{\sigma\sqrt{n}} \sum_{k<i\leq k+j} X_i| \leq 3 \max_{0\leq l\leq m-1;\ 0<j\leq n\delta} |\frac{1}{\sigma\sqrt{n}} \sum_{ln\delta<i\leq ln\delta+j} X_i|.$$

This implies that whenever $m^T(B^n,\delta) > \epsilon$, there exists some $0 \leq l \leq m-1$ such that

$$P[\{\max_{0<j\leq n\delta} |\frac{1}{\sigma\sqrt{n}} \sum_{ln\delta<i\leq ln\delta+j} X_i| > \frac{\epsilon}{3}\}] = 1.$$

Since the number of events of this type is $m = \frac{T}{\delta}$, we therefore have

$$P[\{m^T(B^n,\delta) > \epsilon\}] \leq mP[\{\max_{0<j\leq n\delta} |\frac{1}{\sigma\sqrt{n}} \sum_{1\leq i\leq j} X_i| > \frac{\epsilon}{3}\}]. \tag{55}$$

Setting $S_n = \sum_{i=1}^n X_i$, we see by Kolmogorov's inequality[1] that whenever $\max_{1\leq j\leq n} P[\{|S_n - S_j| > \alpha\}] \leq p < 1$, we have

$$P[\{\max_{1\leq j\leq n} |S_j| > 2\alpha\}] \leq \frac{1}{1-p} P[\{|S_n| > \alpha\}].$$

[1] See the Appendix for the statement

Setting $\alpha = \frac{\epsilon\sigma\sqrt{n}}{6}$ and applying Chebyshev's inequality gives

$$P[\{|\sum_{j+1\leq i\leq n\delta} X_i| > \frac{\epsilon\sigma\sqrt{n}}{6}\}] \leq \frac{6^2\delta n\sigma^2}{\epsilon^2 n\sigma^2} = \frac{36\delta}{\epsilon^2}.$$

In particular, if $36\delta\epsilon^{-2} < 1$, then

$$P[\{\max_{1\leq i\leq j}|\sum_{j+1\leq i\leq n\delta} X_i|] > \frac{\epsilon\sigma\sqrt{n}}{3}\} \leq \frac{1}{1-36\delta\epsilon^{-2}} P[\{|\sum_{1\leq i\leq n\delta} X_i| > \frac{\epsilon\sigma\sqrt{n}}{6}\}]. \tag{56}$$

Lastly, using the Central Limit Theorem, it follows that

$$\begin{aligned}
\limsup_{n\to\infty} P[\{m^T(B_t^n,\delta) > \epsilon\}] &\leq \frac{m}{1-36\delta\epsilon^{-2}} \limsup_{n\to\infty} P[\{|\sum_{1\leq i\leq n\delta} X_i| > \frac{\epsilon\sigma\sqrt{n}}{6}\}] \\
&= \frac{m}{1-36\delta\epsilon^{-2}} 2N(0,1)(\frac{\epsilon}{6\sqrt{\delta}}, \infty) \\
&\leq \frac{2T}{\delta(1-36\delta\epsilon^{-2})} \exp(-\frac{\epsilon^2}{2(6^2\delta)});
\end{aligned}$$

and this goes to 0 as $\delta \to 0$. Therefore, for every $T, \epsilon > 0$ we have that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P[\{m^T(B^n\delta) > \epsilon\}] = 0. \tag{57}$$

We have therefore shown asymptotic equicontinuity holds, and hence Donsker's Theorem follows.

3 Detailed Properties of Brownian Sample Paths

Brownian motion sample paths have very many properties, some desirable, and others not as much. Many of the proofs for these properties require the strong Markov property of Brownian motion, which we will address in the next section of the paper. As a result, we will only state and prove properties of Brownian sample paths that do not require the strong Markov property of Brownian motion. For more properties of Brownian sample paths, see the Exercises at the end of Chapter 2 in Le Gall's text [3]. Before any results can be given, it is necessary to introduce a third definition of Brownian motion.

3.1 Filtrations

Let $B = \{B_t\}_{t\in[0,\infty)}$ be a Brownian motion on the probability space (Ω, Σ, P). Define the σ-algebras Σ_t for $t \in [0,\infty)$ by

$$\Sigma_t := \sigma(\{B_s\}_{s\in[0,t]}).$$

Then

1. whenever $0 \leq s < t$, $\Sigma_s \subset \Sigma_t \subset \Sigma$;
2. for all t, B_t is Σ_t measurable; and
3. for all t, the process $\{B_s - B_t\}_{t\in[s,\infty)}$ is independent of Σ_t.

The first and second points are obvious; the third point follows from the simple Markov property of Brownian motion.

The collection $\{\Sigma_t\}_{t\in[0,\infty)}$ is an example of a *filtration*. To be specific, a filtration on a measurable space (Ω, Σ) is any collection of σ-algebras $\{\Sigma_t\}_{t\in[0,\infty)}$ on Ω that satisfy statements 1-3 as given above. We have just shown that associated to every Brownian motion is a filtration. This filtration will be called the *standard filtration* for B. Filtrations give us a tactical advantage in proving some more advanced properties of Brownian motion. From here on out we will then write a Brownian motion $B = \{B_t\}_{t\in[0,\infty)}$ as $\{(B_t, \Sigma_t)\}_{t\in[0,\infty)}$, where $\{\Sigma_t\}_{t\in[0,\infty)}$ is the standard filtration for B.

For the remainder of this section, let $\{(B_t, \Sigma_t)\}_{t\in[0,\infty)}$ be Brownian motion.

3.2 Blumenthal's Zero-One Law

Define $\Sigma_{0+} := \bigcap_{t>0} \Sigma_t$. We have the following $0-1$ law.

Theorem 6. *(Blumenthal's zero-one law)* For every $A \in \Sigma_{0+}$, either $P[A] = 1$ or $P[A] = 0$.

Proof. Consider any $0 < t_1 < ... < t_n < \infty$, any bounded continuous function $g : \mathbb{R}^k \to \mathbb{R}$, and any fixed $A \in \Sigma_{+0}$. Because g is continuous, B is sample continuous, and $B_0 = 0$ almost surely, we have that

$$E[\chi_A g(B_{t_1}, ..., B_{t_n})] = \lim_{\epsilon\to 0} E[\chi_A g(B_{t_1} - B_\epsilon, ..., B_{t_n} - B_\epsilon)]. \tag{58}$$

By the simple Markov property, whenever $\epsilon < t_1$ the increments $B_{t_n} - B_\epsilon, ..., B_{t_1} - B_\epsilon$ are independent of Σ_ϵ, and therefore are independent of Σ_{+0}. Hence

$$\begin{aligned} E[\chi_A g(B_{t_1}, ..., B_{t_n})] &= \lim_{\epsilon \to 0} P[A]E[g(B_{t_1} - B_\epsilon, ..., B_{t_n} - B_\epsilon)] \\ &= P[A]E[g(B_{t_1}, ..., B_{t_n})]. \end{aligned} \tag{59}$$

This implies that $\sigma(\{B_{t_i}\}_{i=1}^n)$ is independent of Σ_{0+}. Since the t_i were arbitrary, this implies that Σ_{0+} is independent of $\sigma(\{B_t\}_{t\in(0,\infty)})$. Furthermore, for almost every ω, B_0 is the pointwise limit of B_t as $t \to 0$, and therefore $\sigma(\{B_t\}_{t\in(0,\infty)}) = \sigma(\{B_t\}_{t\in[0,\infty)})$. Finally, because $\Sigma_{0+} \subset \sigma(\{B_t\}_{t\in[0,\infty)})$, we see that Σ_{0+} is independent of itself. In particular, this implies that for every $A \in \Sigma_{0+}$, $P[A] = P[A \cap A] = P[A]^2$, so either $P[A] = 0$ or $P[A] = 1$. $\square$

We can use this prove the following proposition.

Proposition 6. 1. For every $\epsilon > 0$, we almost surely have

$$\sup_{0\le t\le\epsilon} B_t > 0 \text{ and } \inf_{0\le t\le\epsilon} B_t < 0. \tag{60}$$

2. For every $a \in \mathbb{R}$, we almost surely have

$$\inf\{t \ge 0 : B_t = a\} < \infty. \tag{61}$$

In particular, we almost surely have

$$\limsup_{t\to\infty} B_t = \infty \text{ and } \liminf_{t\to\infty} B_t = -\infty. \tag{62}$$

Remarks:

1) To see that the function $f_\epsilon(\omega) = \sup_{0\le t\le\epsilon} B_t(\omega)$ is a random variable, we need only notice that by sample continuity, $\sup_{0\le t\le\epsilon} B_t(\omega) = \sup_{0\le t\le\epsilon,\ t\in\mathbb{Q}} B_t(\omega)$. Since the supremum of a countable collection of measurable functions is itself measurable, it follows that $\sup_{0\le t\le\epsilon} B_t$ is measurable, and hence a random variable.

2) Random variables of the form $\tau_a = \inf\{t \ge 0 : B_t = a\}$ for some Brownian motion $\{B_t, \Sigma_t\}_{t\in[0,\infty)}$ are called *hitting times*, and will be discussed in the section about stopping times later on. The first statement of part b of this proposition then states that *hitting times are almost surely finite.*

Proof. 1. Consider any decreasing sequence of positive real numbers $(\epsilon)_{n=1}^\infty$ converging to 0. For each $n = 1, 2, ...,$ define

$$A_n := \{\omega : \sup_{0\le t\le\epsilon_n} B_t(\omega) > 0\}. \tag{63}$$

Then $A_{n+1} \subset A_n$ for every n, and hence $A = \bigcap_{n=1}^\infty A_n$ is Σ_{+0}-measurable. By Blumenthal's zero-one law, either $P[A] = 0$ or $P[A] = 1$. From measure continuity,

$$P[A] = \lim_{n\to\infty} P[A_n], \tag{64}$$

and since $P[A_n] \geq P[\{B_{\epsilon_n} > 0\}] = \frac{1}{2}$ (B_{ϵ_n} is $N(0, \epsilon_n)$ for every n), we see that $P[A] \geq \frac{1}{2}$. It follows that $P[A] = 1$, and in particular that for every $\epsilon > 0$, $\sup_{0 \leq t \leq \epsilon} B_t > 0$ almost surely.

By the symmetry property of Brownian motion, we see that for every $\epsilon > 0$,

$$\begin{aligned} 1 &= P[\{\sup_{0 \leq t \leq \epsilon} (-B_t) > 0\}] \\ &= P[\{-\inf_{0 \leq t \leq \epsilon} B_t > 0\}], \end{aligned} \tag{65}$$

from which the other assertion follows.

2. By the previous part of the proposition and by measure continuity,

$$1 = P[\{\sup_{0 \leq t \leq 1} B_t > 0\}] = \lim_{\epsilon \to 0} P[\{\sup_{0 \leq t \leq 1} B_t > \epsilon\}]. \tag{66}$$

Furthermore, by the scaling invariance property of Brownian motion,

$$P[\{\sup_{0 \leq t \leq 1} B_t > \epsilon\}] = P[\{\sup_{0 \leq t \leq \frac{1}{\epsilon^2}} B_{\epsilon^2 t} > 1\}] = P[\{\sup_{0 \leq t \leq \frac{1}{\epsilon^2}} B_t > 1\}] \tag{67}$$

□

3.3 Zero Set

Define the set Z by

$$Z := \{(t, \omega) \in [0, \infty) \times \Omega : B(t, \omega) = 0\}; \tag{68}$$

and for any fixed $\omega \in \Omega$, define

$$Z(\omega) := \{t \in [0, \infty) : B(t, \omega) = 0\}. \tag{69}$$

We have the following theorem concerning the sets $Z(\omega)$:

Theorem 7. For almost every $\omega \in \Omega$,

1. $Z(\omega)$ has Lebesgue measure 0; and
2. $Z(\omega)$ is closed.

Proof. 1. Noticing that $Z(\omega) \subset (\mathfrak{B}([0, \infty), \Sigma)$, we may apply the Fubini-Tonelli Theorem to get

$$E[\mu(Z(\omega))] = (\mu \times P)(Z) = \int_0^\infty P[\{\omega \in \Omega : B_t(\omega) = 0\}] dt = 0. \tag{70}$$

This implies $\mu(Z(\omega)) = 0$.

2. By definition, for almost every $\omega \in \Omega$, the sample path $\phi : t \to B_t(\omega)$ is continuous; and since $Z(\omega) = \phi^{-1}(0)$, we see that $Z(\omega)$ is closed.

□

3.4 Nowhere Monotonicity

Theorem 8. For almost every $\omega \in \Omega$, the sample path $t \to B_t(\omega)$ is not monotone on any nontrivial interval.

Proof. Fix some $\omega \in \Omega$ such that the sample path $f : t \mapsto B_t(\omega)$ is continuous, and such that Proposition 6(1) holds. Consider the stochastic process $\{B_t - B_s\}_{t\in[s,\infty)}$, for $s > 0$ is rational. By the simple Markov property, this is a Brownian motion. By Proposition 6(1), for every $\epsilon > 0$ we almost surely have

$$\sup_{t\in[s,s+\epsilon]} B_t(\omega) > B_s(\omega) \text{ and } \inf_{t\in[s,s+\epsilon]} B_t(\omega) < B_s(\omega). \tag{71}$$

This implies that f is not monotone on any nontrivial closed interval of $[0,\infty)$. To prove this, suppose on the contrary that f is monotone increasing on some interval $[a,b]$ for $a, b \in \mathbb{Q}, a < b$. Since f is continuous and monotonic increasing on the closed interval $[a,b]$, then $\inf_{t\in[a,b]} B_t(\omega) = B_a(\omega)$. But setting $\epsilon = b-a > 0$, the above inequalities imply $B_a(\omega) < B_a(\omega)$, a contradiction. In particular, this proves that f cannot be monotonic increasing on $[a,b]$. Since a and b were arbitrary rationals with $a < b$, it follows that f cannot be monotone increasing on any nontrivial interval of $[0,\infty)$. □

3.5 Nowhere Differentiability

Theorem 9. The Brownian sample path is nowhere differentiable.

Proof. It suffices to consider a Brownian motion B on a probability space (Ω, Σ, P) with index set $[0,1]$. For $l \in \mathbb{Z}, l \geq 1$, let $A_l = \{s \in [0,1] : \lim_{t\to s} \frac{B_s - B_t}{s-t} \leq l\}$. Our goal is to show that $P[A_l] = 0$ for every l, since this will imply that the collection of all $\omega \in \Omega$ for which the derivative exists has measure zero. For that purpose, assume that $s \in A_l$. I will actually show that there are no three consecutive increments that belong to A_l for any given l. If $s \in [\frac{k}{2^n}, \frac{k-1}{2^n}]$, where $n > 2$, then for every $1 \leq j \leq n$,

$$|B_{\frac{k+j}{2^n}} - B_{\frac{k+(j-1)}{2^n})}| \leq \frac{l(2j+1)}{2^n}. \tag{72}$$

Let $A_{n,k}$ be the event that $|B_{\frac{k+j}{2^n}} - B_{\frac{k+(j-1)}{2^n})}| \leq \frac{l(2j+1)}{2^n}$ holds for $j = 1,2,3$. Then by the scaling invariance of Brownian motion,

$$P[A_{n,k}] \leq P[|B_1| \leq 7l2^{-\frac{n}{2}}]^3 \leq (7l2^{-\frac{n}{2}})^3, \tag{73}$$

where the last inequality follows since B_1 is $N(0,1)$. Therefore,

$$P[\bigcup_{k=1}^{2^n} A_{n,k}] \leq 2^n(7l2^{-\frac{n}{2}})^3 = (7l)^3 2^{-\frac{n}{2}}. \tag{74}$$

Notice that the term on the rightmost hand side is the general term of an absolutely convergent series. Summing over all $n \geq 1$ thus gives $P[\bigcup_{k=1}^{2^n} A_{n,k}] < \infty$. By the Borel-Cantelli lemma,

$$P[\bigcap_{m\geq 1} \bigcup_{n\geq m} (\bigcup_{k\leq 2^n} A_{n,k})] = 0. \tag{75}$$

Since A_l is a subset of $\bigcap_{m\geq 1} \bigcup_{n\geq m}(\bigcup_{k\leq 2^n} A_{n,k})$, it follows that $P[A_l] = 0$[2]. $\square$

3.6 Modulus of Continuity

The following theorem, due to Lévy, gives the exact modulus of continuity of the Brownian sample path.

Theorem 10. For any Brownian motion $B = \{B_t\}_{t\in[0,\infty)}$, we almost surely have

$$\limsup_{0\leq t_1<t_2\leq 1;\ t=t_2-t_1\to 0} \frac{|B_{t_2} - B_{t_1}|}{\sqrt{2t \log \frac{1}{t}}} = 1 \tag{76}$$

Proof. Set $h(t) = \sqrt{2t\log(\frac{1}{t})}$, let $0 < \delta < 1$, and let k, n be nonnegative integers. We have

$$\begin{aligned} P[\{\max_{k\leq n}(B_{k2^{-n}} - B_{(k-1)2^{-n}}) \leq (1-\delta)h(2^{-n})\}] &= (1 - \int_{(1-\delta)h(2^{-n})}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx)^{2^n} \\ &< \exp(-2^n \int_{(1-\delta)h(2^{-n})}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx) \end{aligned}$$

By Lemma 3, there exists a constant $C > 0$ such that

$$\int_{(1-\delta)h(2^{-n})}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx > C\frac{2^n}{\sqrt{n}} \exp(-(1-\delta)^2 \log(2^n)) > 2^{n\delta} \tag{77}$$

whenever n is large enough. Applying the Borel-Cantelli lemma then gives that

$$P[\{\max_{k\leq n} \frac{(B_{k2^{-n}} - B_{(k-1)2^{-n}})}{h(2^{-n})} \geq 1\}] = 1. \tag{78}$$

This proves

$$\limsup_{0\leq t_1<t_2\leq 1;\ t=t_2-t_1\to 0} \frac{|B_{t_2} - B_{t_1}|}{\sqrt{2t \log \frac{1}{t}}} \geq 1. \tag{79}$$

[2] The first result concerning the nowhere differentiability of the Brownian sample path was given by Wiener. The proof given here is due to Dvoretsky et al, and is a much simplified version of Wiener's original proof.

To show the reverse inequality, let $0 < \delta < 1$ and $\epsilon > \frac{1+\delta}{1-\delta} - 1$ be given. We then have

$$P[\{\max_{0<j-i\leq 2^{n\delta};\ 0\leq i<j\leq 2^n} \frac{|B_{j2^{-n}} - B_{i2^{-n}}|}{h((j-i)2^{-n})}\} \geq 1+\epsilon] \leq \sum_{0<j-i\leq 2^{n\delta};\ 0\leq i<j\leq 2^n} 2\int_{(1+\epsilon)\sqrt{2\log(\frac{1}{j-i}2^{-n})}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})dx$$
$$\leq \frac{C}{\sqrt{n}} 2^{n((1+\delta)-(1-\delta)(1+\epsilon)^2)},$$

where $C > 0$ is a constant. Because $\frac{1+\delta}{1-\delta} < (1+\epsilon)^2$, the last expression on the right is the general term of an absolutely convergent series. By the Borel-Cantelli lemma, for $0 \leq i < j \leq 2^n$, $0 < j - i \leq 2^{n\delta}$ and large enough n we have that

$$|B_{j2^{-n}} - B_{i2^{-n}}| < (1+\epsilon)h((j-1)2^{-n}). \tag{80}$$

Choose m large enough so that the inequality above holds whenever $n \geq m$, and choose any $0 \leq s < t \leq 1$ such that $t - s < 2^{-m(1+\delta)}$. Pick some $n \geq m$ so that $2^{-(n+1)(1-\delta)} \leq t - s < 2^{-n(1-\delta)}$, and expand t and s dyadically as follows:

$$s = i2^{-n} - 2^{-p_1} - 2^{-p_2} + ..., \tag{81}$$

where $n < p_1 < p_2 < ...$; and

$$t = j2^{-n} + 2^{-q_1} + 2^{-q_2} + ..., \tag{82}$$

where $n < q_1 < q_2 < ....$ It is easy to see then that $s \leq i2^{-n} < j2^{-n} \leq t$ and that $j - i \leq t2^{-n} < 2^{n\delta}$. By sample continuity of the process B, we see that

$$|B_t - B_s| \leq |B_{i2^{-n}} - B_s| + |B_{j2^{-n}} - B_{i2^{-n}}| + |B_t - B_{j2^{-n}}|$$
$$\leq \sum_{p=n+1}^{\infty} (1+\epsilon)h(2^{-p}) + (1+\epsilon)h((t-s)2^{-n}) + \sum_{p=n+1}^{\infty} (1+\epsilon)h(2^{-p}).$$

Whenever n is large enough, there exists a constant $C > 0$ such that

$$\sum_{p=n+1}^{\infty} h(2^{-p}) \leq Ch(2^{-n}) < \epsilon h(2^{-(n+1)(1-\delta)}); \tag{83}$$

and for small enough t, we have

$$|B_t - B_s| < (1 + 3\epsilon + 2\epsilon^2)h(t). \tag{84}$$

Because ϵ can be chosen however as small as wanted by choosing a sufficiently small δ, it follows that $\limsup_{0\leq t_1<t_2\leq 1;\ t=t_2-t_1\to 0} \frac{|B_{t_2}-B_{t_1}|}{\sqrt{2t\log\frac{1}{t}}} \leq 1$. The theorem now follows. $\square$

4 The Brownian Bridge

Throughout this section we will argue on a probability space (Ω, Σ, P).

Definition 3. A *Brownian bridge* is a sample continuous centered Gaussian process $W = \{W_t\}_{t\in[0,1]}$ such that for all $0 \le s < t \le 1$,

$$C(t, s) = s(1 - t). \tag{85}$$

Brownian bridges exist: for any Brownian motion $B = \{B_t\}_{t\in[0,1]}$, set $W_t = B_t - tB_1$. Then $W = \{W_t\}_{t\in[0,1]}$ is a centered Gaussian process, and for any $0 \le s < t \le 1$ we have

$$\begin{aligned} E[W_tW_s] &= E[(B_t - tB_1)(B_s - sB_1)] \\ &= E[B_tB_s] - sE[B_tB_1] - tE[B_sB_1] + stE[B_1^2] \\ &= s - st - st + st = s(1 - t). \end{aligned}$$

It is important to note that we can 'extract' a Brownian motion from any given Brownian bridge. To see this, let $Y = \{Y_t\}_{t\in[0,1]}$ be a Brownian bridge. Choose any standard Gaussian variable B_1 that is independent of Y_t for every $t \in [0,\infty)$. Then the process $B = \{B_t\}_{t\in[0,1]}$ defined by

$$B_t = Y_t + tB_1 \tag{86}$$

is a Brownian motion. This is easy to see: B is a Gaussian process since Y is and since B_1 is independent of all Y_t. Furthermore, whenever $0 \le s < t \le 1$,

$$\begin{aligned} E[B_sB_t] &= E[Y_sY_t] + sE[Y_sB_1] + tE[Y_tB_1] + stE[B_1^2] \\ &= s(1 - t) + st = s = \min s, t. \end{aligned}$$

In particular, there is a naturally defined bijection between Brownian motions and Brownian bridges. From now on, whenever we consider a Brownian bridge $Y = \{Y_t\}_{t\in[0,1]}$ we will always express it as $Y_t = B_t - tB_1$ for some Brownian motion $B = \{B_t\}_{t\in[0,1]}$. We can actually use this characterization of Brownian bridges in order to compute its probabilities.

To be more specific, let $Y = \{Y_t\}_{t\in[0,1]}$ be the Brownian bridge given by $Y_t = B_t - tB_1$ for the Brownian motion $B = \{B_t\}_{t\in[0,1]}$. For every $\epsilon > 0$, $P[\{|B_1| < \infty\}] \ne 0$; and therefore the conditional law $\mathfrak{L}(B||B_1| < \epsilon)$ is well-defined on $C[0,1]$. Equipping $C[0,1]$ with the usual supremum norm, we get the following theorem.

Theorem 11. On $C[0,1]$, for every $t \in [0,1]$ we have

$$\lim_{\epsilon\to 0^+} \mathfrak{L}(B_t||B_1| < \epsilon) = \mathfrak{L}(Y_t). \tag{87}$$

Proof. Since $E[Y_tB_1] = E[B_tB_1] - tE[B_1^2] = t - t = 0$, the variables Y_t and B_1 are independent. This allows us to write $B_t = Y_t + tB_1$, i.e, we can write B_t as the sum of the Brownian bridge Y and a standard Gaussian variable B_1 that is independent of Y.

Let N_ϵ be the distribution of B_1 given that $|B_1| < \epsilon$, and let Z_ϵ be any random variable that has the distribution N_ϵ and is independent of Y_t. Then

$$\mathfrak{L}(B_t||B_1| < \epsilon) = \mathfrak{L}(Y_t + tZ_\epsilon) = \mathfrak{L}(Y_t) + \mathfrak{L}(tZ_\epsilon), \tag{88}$$

and letting $\epsilon \to 0+$, we see that for all $t \in [0, 1]$,

$$\lim_{\epsilon \to 0+} \mathfrak{L}(B_t||B_1| < \epsilon) = \mathfrak{L}(Y_t). \tag{89}$$

$\square$

4.1 Stopping Times and the Strong Markov Property

Let $\{(B_t, \Sigma_t)\}_{t\in[0,\infty)}$ be a Brownian motion on the probability space (Ω, Σ, P). A *stopping time* for $\{\Sigma_t\}_{t\in[0,\infty)}$ is a random variable $\tau \geq 0$ such that, for every $t \in [0, \infty)$, the event $\{\tau \leq t\}$ is contained in Σ_t.

Stopping times exist. Consider the following three examples:

1. *Constant stopping times*: The constant time variable $\tau(\omega) = c$ is a stopping time, since

$$\{\omega \in \Omega : \tau(\omega) \leq t\} = \begin{cases} \emptyset \in \Sigma_t, & \text{if } t < c; \\ \Omega \in \Sigma_t, & \text{if } t \geq c. \end{cases} \tag{90}$$

2. *Hitting times*: The *hitting time* $\tau_c(\omega) = \inf\{t > 0 : B(t, \omega) = c\}$ is a stopping time, since

$$\{\omega \in \Omega : \tau_c(\omega) \leq t\} = \bigcap_{p<c,\ p\in\mathbb{Q}} \ \bigcup_{q<t,\ q\in\mathbb{Q}} \{\omega \in \Omega : B(q) > p\} \in \Sigma_t \tag{91}$$

by sample continuity.

3. *Dyadic stopping times*: Let τ be any stopping time for the filtration $\{\Sigma_t\}$. Let $\lfloor \tau(\omega) \rfloor$ denote the smallest integer greater than $\tau(\omega)$. The the random variable

$$\tau_n := \frac{\lfloor 2^n \tau \rfloor + 1}{2^n}$$

is a stopping time, called the dyadic stopping time. To see this, note that if $\frac{k}{2^n} \leq \tau < \frac{k+1}{2^n}$, then $\tau_n = \frac{k+1}{2^n}$; whence it follows that for every $t \geq 0$, $\frac{j}{2^n} \leq t < \frac{j+1}{2^n}$ implies

$$\begin{aligned} \{\tau_n \leq t\} = \{\tau_n \leq \frac{j}{2^n}\} &= \{\tau < \frac{j}{2^n}\} \\ &= \bigcup_{q<\frac{j}{2^n},\ q\in\mathbb{Q}} \{\tau \leq q\} \in \Sigma_t. \end{aligned}$$

Note that $\tau_n \geq \tau$ for every $n \in \mathbb{N}$, and that $\tau_n \downarrow \tau$ as $n \to \infty$.

If $\{(X_t, \Sigma_t)\}_{t\in[0,\infty)}$ is a Brownian motion, and τ is a stopping time on $\{\Sigma_t\}_{t\in[0,\infty)}$, then we define

$$\Sigma_\tau := \{A \in \Sigma_\infty : \forall t \in [0,\infty),\ A \cap \{\tau \le t\} \in \Sigma_t\}. \tag{92}$$

It is a quick and easy exercise to show that Σ_τ is in fact a σ-algebra.

We will need the following important fact about stopping times: *any stopping time τ for the filtration $\{\Sigma_t\}_{t\in[0,\infty)}$ is Σ_τ-measurable.*

To see this, first let τ be a stopping time. It suffices to show that the inverse image of every closed interval $[0, s]$ under τ belongs to Σ_τ. So, let $s \in [0,\infty)$ be given. For every $t \in [0,\infty)$, we have

$$\{\tau \le s\} \cap \{\tau \le t\} = \{\tau \le \min\{s,\ t\}\} \in \Sigma_{\min\{s,\ t\}} \subset \Sigma_t. \tag{93}$$

It follows that $\{\tau \le s\} \in \Sigma_\tau$.

We can now move on to our main result concerning stopping times, namely, the *strong Markov property of Brownian motion*. Colloquially, this result states that *Brownian motion begins afresh at stopping times*. We will use this result frequently for the rest of the paper, as it will be an essential ingredient in proving many of the different bounds for probabilities concerning Brownian bridges.

Theorem 12. (*Strong Markov Property*) On the event $\{\tau < \infty\}$, the process $B' = \{B'_t\}$ defined by

$$B'_t = B_{\tau+t} - B_\tau$$

is Brownian motion, independent of Σ_τ.

Proof. Consider the dyadic stopping times τ_n as defined before. Since B is sample continuous and $\tau_n \downarrow \tau$, it follows that $B_{\tau_n} \to B_\tau$ almost surely. Since B' is sample continuous by definition, it suffices to check that B' has the same finite dimensional distributions as pre-Brownian motion. To do this, let $k \in \mathbb{Z}_{\ge 1}$ be given. Choose any $0 \le t_1 < t_2 < ... < t_k$, any bounded continuous function $F : \mathbb{R}^k \to \mathbb{R}$, and fix $A \in \Sigma_\tau$. Define

$$G(t) := F(B_{\tau+t_1} - B_\tau, ..., B_{\tau+t_k} - B_\tau). \tag{94}$$

Then G is bounded and continuous on $\mathbb{R}^k$. Furthermore, whenever $j \ge 0$, the event

$$A \bigcap \{\tau_n = \frac{j}{2^n}\} = B \bigcap \{\frac{j-1}{2^n} \le \tau < \frac{j}{2^n}\} \in \Sigma_{\frac{j}{2^n}}$$

is independent of $G(\frac{l}{2^n})$, which in turn implies

$$\begin{aligned} E[G(\tau_n)\chi_A\chi_{\{\tau<\infty\}}] &= \sum_{j\ge 0} E[G(\frac{j}{2^n})\chi_{B\cap\{\tau_n=\frac{k}{2^n}\}}] \\ &= \sum_{j\ge 0} E[G(\frac{j}{2^n})]E[\chi_{B\cap\{\tau_n=\frac{j}{2^n}\}}] \\ &= E[G(0)]\sum_{j\ge 0} E[\chi_{B\cap\{\tau_n=\frac{j}{2^n}\}}] \\ &= E[G(0)]E[\chi_B\chi_{\{\tau<\infty\}}]. \end{aligned}$$

Here, we used the independence as stated above in the second line, and the homogeneity of Brownian motion in the third line in order to write $E[G(\frac{j}{2^n})] = E[G(0)]$. By continuity of G, $G(\tau_n) \to G(\tau)$, and hence

$$E[G(\tau)\chi_B\chi_{\{\tau<\infty\}}] = \lim_{n\to\infty} E[G(\tau_n)\chi_B\chi_{\{\tau<\infty\}}] = E[G(0)]E[\chi_B\chi_{\{\tau<\infty\}}]. \quad (95)$$

In particular, since $B \in \Sigma_t$ was arbitrary, this shows that the the increments $(B'_{t_1}, ..., B'_{t_k})$ are independent of the σ-algebra Σ_t. Furthermore, it shows that, on the event $\{\tau < \infty\}$, the distribution of $(B'_{t_1}, ..., B'_{t_k})$ is the same as the distribution of $(B_{t_1}, ..., B_{t_n})$, so that B' has the same finite dimensional distributions as pre-Brownian motion, and hence is itself Brownian motion. □

4.2 Reflection Principle for Brownian Motion

Let $\{(B_t, \Sigma_t)\}_{t\in[0,\infty)}$ be a Brownian motion on (Ω, Σ, P). We have the following "reflection principle":

Theorem 13. For any $t \in [0,1]$, set $S_t = \sup_{s\in[0,t]} B_s$. Then for any $a \in [0,\infty)$ and any $b \in (-\infty, a]$, we have

$$P[\{S_t \geq a \text{ and } B_t \leq b\}] = P[\{B_t \geq 2b - a\}] \quad (96)$$

In particular, S_t and $|B_t|$ have the same distribution.

Remark: This result will be referred to as the *reflection principle for Brownian motion.*

Proof. Consider the hitting time $\tau_a = \inf\{t > 0 : B_t = a\}$. We have that $\{\sup_{0\leq s\leq t} B_s \geq a\} = \{\inf\{s > 0 : B_s = a\} \leq t\}$; and, from Proposition 6(2), we have that $\tau_a < \infty$ almost surely. Furthermore, using the notation from the statement of the strong Markov property, we have by definition of τ_a that $B'_{t-\tau_a}(\omega) = B_t(\omega) - B_{\tau_a(\omega)}(\omega) = B_t(\omega) - a$. Thus

$$\begin{aligned} P[\{S_t \geq a,\ B_t \leq b\}] &= P[\{\tau_a \leq t,\ B_t \leq b\}] \\ &= P[\{\tau_a \leq t,\ B^{(\tau_a)}_{t-\tau_a} \leq b - a\}]. \end{aligned}$$

By the Strong Markov property, B' is a Brownian motion and is independent of Σ_{τ_a}, and therefore is independent of τ_a. Because both B' and $-B'$ have the same law, the pairs (τ_a, B') and $(\tau_a, -B')$ also have the same law. Setting $T = \{(s,f) \in \mathbb{R}^+ \times C(\mathbb{R}^+, \mathbb{R}) : s \leq t,\ f(t-s) \leq b - a\}$, we see that $P[\{\tau_a \leq t,\ B^{(\tau_a)}_{t-\tau_a} \leq b - a\}] = P[\{\exists t \in [0,\infty) : (\tau_a, B'_t) \in T\}]$; and therefore

$$\begin{aligned} P[\{\tau_a \leq t,\ B^{(\tau_a)}_{t-\tau_a} \leq b - a\}] &= P[\{\exists t \in [0,\infty) : (\tau_a, B'_t) \in T\}] \\ &= P[\{\tau_a \leq t,\ -B^{(\tau_a)}_{t-\tau_a} \leq b - a\}] \\ &= P[\{\tau_a \leq t,\ B_t - a \geq a - b\}] \\ &= P[\{\tau_a \leq t,\ B_t \geq 2a - b\}] \\ &= P[\{S_t \geq a, B_t \geq 2a - b\}] \\ &= P[\{B_t \geq 2a - b\}], \end{aligned} \quad (97)$$

where the last equality follows since $B_t \geq 2a - b$ implies $S_t \geq a$.

Furthermore, we have

$$\begin{aligned} P[S_t \geq a\}] &= P[\{S_t \geq a,\ B_t \geq a\}] + P[\{S_t \geq a,\ B_t \leq a\}] \\ &= P[\{B_t \geq a\}] + P[\{B_t \geq a\}] = 2P[\{B_t \geq a\}] \\ &= P[\{|B_t| \geq a\}], \end{aligned} \tag{98}$$

which shows that S_t and $|B_t|$ have the same law. □

In particular, this theorem implies that the law of the pair (S_t, B_t) has density

$$\rho(x,y) = \frac{2(2x-y)}{\sqrt{2\pi t^3}} \exp\left(-\frac{(2x-y)^2}{2t}\right)\chi_{\{a>0,b<a\}}. \tag{99}$$

Moreover, for any $a > 0$, we may use the above theorem and the scaling invariance of Brownian motion to see that

$$\begin{aligned} P[\{\tau_a \leq t\}] &= P[\{S_t \geq a\}] \\ &= P[\{|B_t| \geq a\}] \\ &= P[\{B_t^2 \geq a^2\}] \\ &= P[\{B_t^2 \geq a^2\}] \\ &= P[\{tB_1^2 \geq a^2\}] \\ &= P[\{\frac{a^2}{B_1^2} \leq t\}]. \end{aligned} \tag{100}$$

Using the fact that B_1 is a centered Gaussian variable with variance 1, we can easily calculate the density function of $\frac{a^2}{B(1)^2}$ as follows. Consider the random variables $X = B_1$, $Y = \frac{X^2}{a^2}$, and $Z = \frac{1}{Y}$ with their respective densities $f(x)$, $g(y)$, and $h(z)$. Then a simple computation shows that $g(y) = \frac{a}{2\sqrt{t}} f(a\sqrt{y})$; and that $h(z) = \frac{1}{z^2} g(\frac{1}{z}) = \frac{a}{2z\sqrt{z}} f(\frac{a}{\sqrt{z}})$. Substituting in $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$ and remembering that we are conditioned on the fact that $t > 0$ shows the density of $\frac{a^2}{B_1^2}$, and thus the density of τ_a, is given by

$$h(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right)\chi_{\{t>0\}}. \tag{101}$$

.

4.3 Reflection Principles for Brownian Bridges

Assume that $Y = \{Y_t\}_{t\in[0,1]}$ is a Brownian bridge.

Theorem 14. For any $b \in \mathbb{R}$,

$$P[\sup_{t\in[0,1]} Y_t \geq b] = \exp(-2b^2). \tag{102}$$

Proof. Define $B_t = Y_t + tB_1$, where B_1 is $N(0,1)$ and is independent of Y_t. Then $B = \{B_t\}_{t\in[0,1]}$ is a Brownian motion. Since $E[Y_tB_1] = E[B_tB_1] - tE[B_1^2] = t - t = 0$, Y_t and B_1 are independent. Setting $Q = \{\exists t \in [0,1] : Y_t = b\}$, we may write

$$\begin{aligned} Q &= \frac{P[\{\exists t \in [0,1] : Y_t = b; |B_1| < \epsilon\}]}{P[\{|B_1| < \epsilon\}]} \\ &= \frac{P[\{\exists t \in [0,1] : B_t = b + tB_1; |B_1| < \epsilon\}]}{P[\{|B_1| < \epsilon\}]}. \end{aligned} \tag{103}$$

Letting Q' be the numerator in this expression, we have the following upper and lower bounds for Q':

$$P[\{\exists t \in [0,1] : B_t \geq b + \epsilon; |B_1| < \epsilon\}] \leq Q'; \tag{104}$$

$$Q' \leq P[\{\exists t \in [0,1] : B_t \geq b - \epsilon; |B_1| < \epsilon\}]. \tag{105}$$

We will examine both the upper bound and the lower bound. We will start with the lower bound. Consider the hitting time $\tau = \inf\{t > 0 : B_t = b + \epsilon\}$. Then $\tau < \infty$ almost surely by Proposition 6(2). Moreover, for almost every ω we have $B_{t-\tau}^{(\tau)}(\omega) = B_t(\omega) - B_\tau(\omega) = B_t(\omega) - (b + \epsilon)$, and therefore

$$\begin{aligned} Q' &= P[\{\tau \leq 1; B_1 - B_\tau \in (-b - 2\epsilon, -b)\}] \\ &= P[\{\tau \leq 1; B_{1-\tau}^{(\tau)} \in (-b - 2\epsilon, -b)\}]. \end{aligned} \tag{106}$$

Using the strong Markov Property and the symmetry of Brownian motion, we see that

$$\begin{aligned} Q &= P[\{\tau \leq 1; -B_{1-\tau}^{(\tau)} \in (-b - 2\epsilon, -b)\}] \\ &= P[\{\tau \leq 1; B_{1-\tau}^{(\tau)} \in (b, b + 2\epsilon)\}] \\ &= P[\{\tau \leq 1; B_1 \in (2b + \epsilon, 2b + 3\epsilon)\}] \\ &= P[\{B_1 \in (2b + \epsilon, 2b + 3\epsilon)\}], \end{aligned} \tag{107}$$

where in the last line we used the fact that $B_1 \in (2b + \epsilon, 2b + 3\epsilon)$ implies that $\tau \leq 1$ whenever $b > 0$. Thus

$$P[\{\exists t \in [0,1] : Y_t = b\}] \geq \frac{P[\{B_1 \in (2b + \epsilon, 2b + 3\epsilon)\}]}{P[\{B_1 \in (-\epsilon, \epsilon)\}]}. \tag{108}$$

Since

$$\frac{P[\{B_1 \in (2b + \epsilon, 2b + 3\epsilon)\}]}{P[\{B_1 \in (-\epsilon, \epsilon)\}]} = \frac{\int_{2b+\epsilon}^{2b+3\epsilon} \exp(-\frac{x^2}{2})dx}{\int_{-\epsilon}^{\epsilon} \exp(-\frac{x^2}{2})dx}, \tag{109}$$

we let $\epsilon \to 0+$ to see that

$$P[\{\exists t \in [0,1] : Y_t = b\}] \geq e^{-2b^2}. \tag{110}$$

The upper bound my be examined similarly using the starting time $\tau = \tau(\omega) = \inf\{t > 0 : B_t = b - \epsilon\}$, which will yield the inequality

$$P[\{\exists t \in [0,1] : Y_t = b\}] \leq \frac{P[\{B_1 \in (2b - 3\epsilon, 2b - \epsilon)\}]}{P[\{B_1 \in (-\epsilon, \epsilon)\}]}.$$

Since

$$\frac{P[\{B_1 \in (2b-3\epsilon, 2b-\epsilon)\}]}{P[\{B_1 \in (-\epsilon, \epsilon)\}]} = \frac{\int_{2b-3\epsilon}^{2b-\epsilon} \exp(-\frac{x^2}{2})dx}{\int_{-\epsilon}^{\epsilon} \exp(-\frac{x^2}{2})dx},$$

we let $\epsilon \to 0+$ to see that

$$P[\{\exists t \in [0,1] : Y_t = b\}] \leq e^{-2b^2}. \tag{111}$$

Equality follows by combining the inequalities for the upper and lower bounds. $\square$

4.3.1 The Brownian Bridge as the Limit of the Empirical Process, and the Kolmogorov-Smirnov Distribution

Given any sequence $(X_n)_{n=1}^{\infty}$ of independent identically-distributed random variables with (perhaps unknown) cumulative distribution function $F(t) := P[\{X_1 \leq t\}]$, an *empirical cumulative distribution function* for F is a function of the form

$$F_n(t) := \frac{1}{n}\sum_{i=1}^{n} \chi_{\{X_i \leq t\}}. \tag{112}$$

First consider any sequence $(X_n)_{n=1}^{\infty}$ of independent identically distributed uniform random variables on $[0,1]$. Let $F_n(t) := n^{-1}\sum_{i=1}^{n} \chi_{\{X_i \leq t\}}$ be an empirical cumulative distribution function for F. By the Strong Law of Large Numbers, for any $t \in [0,1]$, the sequence $(F_n(t))_{n=1}^{\infty}$ converges to the cumulative distribution function $F(t) := P[\{X_1 \leq t\}] = t$ almost surely. Furthermore, by the Central Limit Theorem, the *empirical process* $\{X_t^n\}_{t\in[0,1]}$ defined by

$$X_t^n := \sqrt{n}(n^{-1}\sum_{i=1}^{n} \chi_{\{X_i \leq t\}} - t) \tag{113}$$

converges in distribution to $N(0, t(1-t))$. Additionally, the process $\{X_t^n\}_{t\in[0,1]}$ has covariance function $C(s,t) = s(1-t)$, since $s \leq t$ implies

$$E[X_s^n X_t^n] = E[(\chi_{\{X_1 \leq s\}} - s)(\chi_{\{X_1 \leq t\}} - t)] = s - ts - st + st = s(1-t). \tag{114}$$

This is the same covariance as the Brownian bridge $B = \{B_t\}_{t\in[0,1]}$. By the Multivariate Central Limit Theorem, the finite dimensional distributions of the empirical process converges to the finite dimensional distributions of the Brownian bridge, i.e, for every finite subset $F \subset [0,1]$,

$$\mathfrak{L}((X_t^n)_{t\in F}) \to \mathfrak{L}((B_t)_{t\in F}). \tag{115}$$

From the previous discussion about Brownian motion being the limit of distributions, this identifies the law of Brownian bridge B as the unique possible limit of $\{\mathfrak{L}(X_t^n)\}_{n=1}^{\infty}$. In fact, we have the weak convergence $\mathfrak{L}(X_t^n) \to \mathfrak{L}(B_t)$. However, a much stronger statement actually holds; namely, *for every* $\phi \in (C[0,1], ||\cdot||_{\infty})$,

$\{\mathfrak{L}(\phi(X_t^n))\}_{n=1}^{\infty}$ *converges weakly to* $\mathfrak{L}(\phi(B_t))$. See [1], Section 23 for more details regarding the proof of this weak convergence.

We would like to give an application of this result. In order to do this, consider some independent identically distributed random variables $(X_i)_{i=1}^{\infty}$ with continuous cumulative distribution function $F(t) = P[\{X_1 \le t\}]$, and consider the empirical cumulative distribution functions $F_n(t) = n^{-1}\sum_{i=1}^{n}\chi_{\{X_i \le t\}}$. Because F is continuous, the range of F is the whole interval $[0,1]$; and because $(F(X_i))_{i=1}^{\infty}$ are independent identically distributed uniform random variables, we have

$$\begin{aligned}
\sup_{t\in\mathbb{R}} \sqrt{n}|F_n(t) - F(t)| &= \sup_{t\in\mathbb{R}} |n^{-1}\sum_{i=1}^{n}\chi_{\{X_i\le t\}} - F(t)| \\
&= \sup_{t\in\mathbb{R}} \sqrt{n}|n^{-1}\sum_{i=1}^{n}\chi_{\{F(X_i)\le F(t)\}} - F(t)| \qquad (116)\\
&= \sup_{t\in[0,1]} \sqrt{n}|n^{-1}\sum_{i=1}^{n}\chi_{\{F(X_i)\le t\}} - t|,
\end{aligned}$$

The last expression here is the distribution of $\sup_{t\in[0,1]}|X_t^n|$. By the above result, we have the weak convergence $\{\mathfrak{L}(\sup_{t\in[0,1]}|X_t^n|)\}_{n=1}^{\infty} \to \mathfrak{L}\sup_{t\in[0,1]}|B_t|$, where $\{B_t\}_{t\in[0,1]}$ is the Brownian bridge. The distribution $\mathfrak{L}\sup_{t\in[0,1]}|B_t|$ is called the *Kolmogorov-Smirnov distribution*, and we can actually give an explicit expression for it.

Theorem 15. *(Kolmogorov-Smirnov)* For any Brownian bridge $Y = \{Y_t\}_{t\in[0,1]}$ and any $b > 0$,

$$P[\sup_{t\in[0,1]}|Y_t| \ge b] = 2\sum_{n=1}^{\infty}(-1)^{n-1}\exp(-2n^2b^2). \qquad (117)$$

Proof. Let $A_n = \{\omega \in \Omega : \exists 0 < t_1 < \ldots < t_n \le 1 \ni W_{t_j} = (-1)^{j-1}b\}$, and let $\tau_b = \inf\{t \in [0,1] : Y_t = b\}$ and $\tau_{-b} = \{t \in [0,1] : Y_t = -b\}$ be the hitting times of b and b, respectively. Set $Q_n = P[A_n \cap \{\tau_b < \tau_{-b}\}]$. Because Q_{n+1} is unchanged by interchanging Y_t and $-Y_t$, we see that $Q_n = P[A_n] - Q_{n+1}$. By Theorem 14, $P[A_1] = \exp(-2b^2)$.

We now need to make the following observation. The probability that $|Y_t|$ reaches b is the same as the probability that Y_t reaches b or $-b$, which is twice the probability that Y_t reaches b minus the probability that Y_t reaches both b and $-b$. Now consider the Brownian motion $B = \{B_t\}_{t\in[0,1]}$ defined by $B_t = Y_t + tB_1$, where B_1 is $N(0,1)$ and is independent of Y_t. By Theorem 11, we need only consider the probability that B_t reaches b, then reaches $-b$, and then $|B_t| < \epsilon$. This is the same as the probability that B_t reaches b, then reaches $3b$, and then satisfies $|B_t - 4b| < \epsilon$. But whenever $\epsilon < b$, this inequality implies that X_t has reached both b and $3b$, and therefore the probability being considered is just the probability that $|B_1 - 4b| < \epsilon$. In particular, we have that

$$\begin{aligned} P[A_2] =& \lim_{\epsilon\to\infty} P[\{|B_1 - 4b| < \epsilon\}] \\ =& \exp(-\frac{1}{2}(4b)^2) = \exp(-8b^2) = \exp(-2(2)^2 b^2). \end{aligned} \tag{118}$$

Now consider A_3. Again by Theorem 11, we need only consider the probability that B_t hits b, then hits $-b$, then hits b, and then satisfies $|B_1| < \epsilon$. This is the probability that B_t hits b, then hits $3b$, then hits $5b$, and then satisfies $|B_1 - 6b| < \epsilon$. The event that B_t hits b, then hits $3b$, and finally hits $5b$ contains the event that $|B_1 - 6b| < \epsilon$, and therefore the probability being considered is simply the probability of the event $\{|B_1 - 6b| < \epsilon\}$. This gives

$$\begin{aligned} P[A_3] =& \lim_{\epsilon\to\infty} P[\{|B_1 - 6b| < \epsilon\}] \\ =& \exp(-\frac{36b^2}{2}) = \exp(-18b^2) = \exp(-2(3)^2 b^2). \end{aligned} \tag{119}$$

In a more general setting considering A_n, we can apply Theorem 11 and the reflection principle to the Brownian motion B and to see that $P[A_n]$ is the limit as $\epsilon \to 0$ of the probability that we first hit b, and then make $n-1$ successive increasing jumps of size $2b$, and then satisfy $|B_0 - (2+2(n-1))b| = |B_0 - 2nb| < \epsilon$. The sequence of hits is contained in the event that $|B_0 - 2nb| < \epsilon$, and so the desired probability is simply the latter probability. Letting $\epsilon \to 0$ gives that $P[A_n] = \exp(-2n^2b^2)$. This implies that $Q_n \to 0$ as $n \to \infty$; and therefore that

$$\begin{aligned} P[\sup_{t\in[0,1]} |W_t| \geq b] &= 2Q_1 = 2[P[A_1] - Q_2] \\ &= 2[P[A_1] - P[A_2] + Q_3] \\ &= \cdots \\ &= 2[\sum_{n=1}^{\infty} (-1)^{n-1} P[A_n]] \\ &= 2[\sum_{n=1}^{\infty} (-1)^{n-1} \exp(-2n^2b^2)]. \end{aligned} \tag{120}$$

□

Remark: This distribution is often given by.

$$P[\sup_{t\in[0,1]} |Y_t| < b] = 1 - 2\sum_{n=1}^{\infty} (-1)^{n-1} \exp(-2n^2b^2). \tag{121}$$

4.3.2 More Bounds for Probabilities Regarding the Brownian Bridge

We also can give an explicit expression for the probability that a Brownian bridge remains between the levels $-a$ or b for any given $a, b > 0$.

Theorem 16. (*Two Sided Boundary*) For any Brownian bridge $Y = \{Y_t\}_{t\in[0,1]}$ and any $a, b > 0$,

$$P[\exists t : Y_t = -a \text{ or } b] = \sum_{n=0}^{\infty}(e^{-2(na+(n+1)b)^2} + e^{-2((n+1)a+nb)^2}) - 2\sum_{n=1}^{\infty} e^{-2n^2(a+b)^2}. \tag{122}$$

Proof. Let $A_n = \{\exists t_1 < ... < t_n : Y_{t_j} = -a, \text{ j odd}; Y_{t_i} = b, \text{ i even}\}$, and let $B_n = \{\exists t_1 < ... < t_n : Y_{t_j} = -a, \text{ j even}; Y_{t_i} = b, \text{ i odd}\}$. Consider the hitting times $\tau_{-a} = \inf\{t : Y_t = -a\}$ and $\tau_b = \inf\{t : t_t = b\}$. Then we have

$$P[\exists t : Y_t = -a \text{ or } b] = P[\exists t : Y_t = -a;\ \tau_{-a} < \tau_b] + P[\exists t : Y_t = b;\ \tau_b < \tau_{-a}]. \tag{123}$$

Similarly to the previous theorem, we can see that

$$P[B_n; \tau_b < \tau_{-a}] = P[B_n] - P[B_n; \tau_{-a} < \tau_b] = P[B_n] - P[A_{n+1}; \tau_{-a} < \tau_b]; \tag{124}$$

and

$$P[A_n; \tau_{-a} < \tau_b] = P[A_n] - P[A_n; \tau_b < \tau_{-a}] = P[A_n] - P[B_{n+1}; \tau_b < \tau_{-a}]. \tag{125}$$

By induction, we therefore have

$$P[\exists t : Y_t = -a \text{ or } b] = \sum_{n=1}^{\infty}(-1)^{n-1}(P[A_n] + P[B_n]). \tag{126}$$

It only remains to calculate the probabilities $P[A_n]$ and $P[B_n]$ and substitute them into this expression. This is easy, and can be done using the previous reflection principles to get

$$P[A_{2n}] = P[B_{2n}] = e^{-2n^2(a+b)^2}; \tag{127}$$

$$P[B_{2n+1}] = e^{-2(na+(n+1)b)^2}; \text{ and} \tag{128}$$

$$P[A_{2n+1}] = e^{-2((n+1)a+nb)^2}. \tag{129}$$

□

Our final theorem gives an exact expression for the cumulative distribution function for the random variable $Y' = \sup_{t\in[0,1]} Y_t - \inf_{t\in[0,1]} Y_t$, where $Y = \{Y_t\}_{t\in[0,1]}$ is a Brownian bridge.

Theorem 17. For any Brownian bridge $Y = \{Y_t\}_{t\in[0,\infty)}$ and any $h > 0$,

$$P[\sup_{t\in[0,1]} Y_t - \inf_{t\in[0,1]} Y_t \le h] = 1 - \sum_{n=1}^{\infty}(8n^2h^2 - 2)e^{-2n^2h^2}. \tag{130}$$

Proof. Let $X = -\inf_{t\in[0,1]} Y_t$ and let $Z = \sup_{t\in[0,1]} Y_t$, and let $F(a,b)$ be the joint cumulative distribution function of (X,Z). This function we already know, since

$$F(a,b) = P[X < a; Z < b] = P[-a < \inf_{t\in[0,1]} Y_t; \sup_{t\in[0,1]} Y_t < b] = 1 - P[\exists t : Y_t = -a \text{ or } b],$$

where the last expression can be calculated using Theorem 16. Letting $f(a,b) = \frac{\partial^2 F}{\partial a \partial b}$ be the joint distribution function of X and Z, we can calculate the cumulative distribution function of $X + Z$ as follows:

$$P[X + Z \leq h] = \int_0^h \int_0^{h-a} f(a,b)dbda. \tag{131}$$

Evaluating the inner integral gives

$$\int_0^{h-a} f(a,b)db = \frac{\partial F}{\partial a}(a, h-a) - \frac{\partial F}{\partial a}(a, 0). \tag{132}$$

Differentiating the expression in Theorem 16, we can calculate $\frac{\partial F}{\partial a}$ as follows:

$$\begin{aligned} \frac{\partial F}{\partial a}(a,b) = &\sum_{n=0}^{\infty} 4n(na + (n+1)b)e^{-2(na+(n+1)b)^2} \\ &+ \sum_{n=0}^{\infty} 4(n+1)((n+1)a + nb)e^{-2((n+1)a+nb)^2} \\ &- \sum_{n=1}^{\infty} 8n^2(a+b)e^{-2n^2(a+b)^2}. \end{aligned} \tag{133}$$

Setting $b = h - a$ and $b = 0$ above and subtracting the latter from the former gives

$$\begin{aligned} \int_0^{h-a} f(a,b)db = &\sum_{n=0}^{\infty} 4n((n+1)h - a)e^{-2((n+1)h-a)^2} \\ &+ \sum_{n=0}^{\infty} 4(n+1)(nh + a)e^{-2(nh+a)^2} \\ &- \sum_{n=1}^{\infty} 8n^2h^2e^{-2n^2h^2}. \end{aligned} \tag{134}$$

Finally, integrating this expression over $a \in [0,h]$ gives us

$$\begin{aligned} P[X + Z \leq h] &= \sum_{n=0}^{\infty}(2n+1)(e^{-2n^2h^2} - e^{-2(n+1)^2h^2}) - \sum_{n=1}^{\infty} 8n^2h^2e^{-2n^2h^2} \\ &= 1 - \sum_{n=1}^{\infty}(8n^2h^2 - 2)e^{-2n^2h^2} \end{aligned} \tag{135}$$

as wanted. ◻

5 Appendix

5.1 Kolmogorov's Inequality

Lemma 6. *(Kolmogorov's Inequality)* Let $(X_n)_{n=1}^{\infty}$ be any sequence of independent random variables, and set $S_n := \sum_{i=1}^{n} X_i$. If, for every $j \leq n$ we have that

$$P[\{|S_n - S_j| \geq a\}] \leq p < 1,$$

then whenever $x > a$,

$$P[\{\max_{1\leq j\leq n} |S_j| \geq x\}] \leq \frac{1}{1-p} P[\{|S_n| > x - a\}]. \tag{136}$$

I will not give the proof of this inequality here, since it requires a bit of work. Refer to [1], Theorem 16 for the proof.

5.2 The Central Limit Theorem and the Law of Large Numbers

Theorem 18. *((Strong) Law of Large Numbers)* Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of independent, identically-distributed random variables on the same probability space (Ω, Σ, P), and assume that $E[|X_1|] < \infty$. Then almost surely,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to E[X_1] \tag{137}$$

as $n \to \infty$.

Theorem 19. *(Central Limit Theorem)* Let $\{X_n\}_{n=1}^{\infty}$ be any independent, identically-distributed copies of a random variable X on the probability space (Ω, Σ, P), each with mean 0 and finite variance σ^2. Then

$$S_n = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n} X_i \tag{138}$$

converges in distribution to $N(0, 1)$ as $n \to \infty$.

Theorem 20. *(Multivariate Central Limit Theorem)* Let $X = (x_1, ..., x_n)$ be any random vector on the probability space (Ω, Σ, P) with mean 0 and covariance matrix C, and let $\{X_n\}_{n=1}^{\infty}$ be any sequence of independent, identically-distributed copies of X. Then

$$S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - E[X_i]) \tag{139}$$

converges in distribution to $N(0, C)$ as $n \to \infty$.

References

[1] Theory of Probability Notes.
https://goo.gl/lDXzBz

[2] H. P. McKean, Jr. *Stochastic Integrals.* The Rockefeller University. Academic Press, Inc. New York, New York, 1969. pg 1-19.

[3] Jean-Francois Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus.* Springer-Verlag, Berlin, Germany. 2013. pg. 1-40.

[4] Gaussian Distributions with Application, MAT 477, Fall 2016.
https://goo.gl/b5dJN2