

Bowdoin College

Bowdoin Digital Commons

Honors Projects

Student Scholarship and Creative Work

2019

An alternative almost sure construction of Gaussian stochastic processes in the $L_2([0,1])$ space

Kevin Chen
kchen.8298@gmail.com

Follow this and additional works at: <https://digitalcommons.bowdoin.edu/honorsprojects>



Part of the [Analysis Commons](#), [Geometry and Topology Commons](#), [Harmonic Analysis and Representation Commons](#), [Other Mathematics Commons](#), and the [Set Theory Commons](#)

Recommended Citation

Chen, Kevin, "An alternative almost sure construction of Gaussian stochastic processes in the $L_2([0,1])$ space" (2019). *Honors Projects*. 126.

<https://digitalcommons.bowdoin.edu/honorsprojects/126>

This Open Access Thesis is brought to you for free and open access by the Student Scholarship and Creative Work at Bowdoin Digital Commons. It has been accepted for inclusion in Honors Projects by an authorized administrator of Bowdoin Digital Commons. For more information, please contact mdoyle@bowdoin.edu.

An alternative almost sure construction
of Gaussian stochastic processes
in the $L^2([0, 1])$ space

An Honors Paper for the Department of Mathematics
By Kevin Chen

Bowdoin College, 2019
©2019 Kevin Chen

Acknowledgements

I would like to begin by thanking my advisor, Patricia Garmirian, for all her guidance and support. Working with her has not only been a pleasure, but also a deeply enriching learning experience. I believe that I have grown the most as a mathematician during my time spent with her.

The mathematics department at Bowdoin as a whole has been truly outstanding. I would like to particularly thank William Barker, Behrang Forghani and Adam Levy for their feedback on my thesis. I would also like to thank Michael King and Jack O'Brien for being the most resourceful academic advisors that I have had throughout my career as a student. Finally, I would like to thank Thomas Pietraho for introducing me to the field of mathematical analysis.

To continue with the trend of thanking my past instructors, I would like to thank William Fillbach and Leo Lopez. I am lucky to have had them both as instructors during my younger years.

Finally, these acknowledgements would not be complete without extending the utmost thanks to my parents. I would not be where I am today if it were not for them. I would like to particularly thank my dad for introducing mathematics into my life and for supporting me along my mathematical journey.

Contents

1	Preliminary background	4
1.1	Measure theoretic probability and stochastic processes.	4
1.2	Hilbert space theory.	10
1.3	Convergence.	16
2	Strong convergence of functions in the $L^2([0, 1])$ space via Haar basis	18
2.1	L^p space.	19
2.2	L^2 space.	20
2.3	Haar basis.	21
3	Brownian motion	27
3.1	Brownian motion as a sample path in $L^2([0, 1])$	28
3.2	Haar basis expansion of Brownian motion.	29
4	Ornstein–Uhlenbeck process	32
4.1	Ornstein–Uhlenbeck process as a sample path in $L^2([0, 1])$	33
4.2	Haar basis expansion of the Ornstein–Uhlenbeck process.	34
5	Final remarks	36
	Appendices	38
A	Zorn’s Lemma	38
B	Fatou’s Lemma	39
C	Fubini’s Theorem	40
D	Brownian motion simulation	43

Introduction

A stochastic processes is a sequence of random variables $\{X_t \mid t \geq 0\}$ defined on a common probability space (Ω, \mathcal{F}, P) , indexed by time t . Stochastic processes seek to model random processes over both discrete and continuous times. The study of stochastic processes uses techniques from probability theory, topology, calculus and mathematical analysis. Stochastic processes are currently of high interest for their applications in the financial markets, natural sciences and technology fields.

Brownian motion is an elementary and commonly studied stochastic process. Brownian motion was first discovered by Scottish botanist Robert Brown in 1827 while observing the movement of pollen grain in water. Traditionally, Brownian motion has been constructed as a weak distributional limit of random walks. Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d.) Gaussian random variables with mean 0 and variance 1. Define a stochastic process $\{B_t^n\}$:

$$B_t^n = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nt \rfloor} X_j$$

From the Central Limit Theorem, $\lim_{n \rightarrow \infty} B_t^n$ converges in distribution to Brownian motion. However, we would like an almost sure construction of Brownian motion. Other constructions of Brownian motion include French mathematician Paul Lévy's almost sure construction that linearly interpolates Brownian motion over the dyadic rationals.

The motivation for this paper is to provide an alternative construction for stochastic processes that is almost sure. The focus will be on Gaussian stochastic processes, specifically Brownian motion and the Ornstein–Uhlenbeck process.

Our construction starts with the nice property of Hilbert spaces that any functions in the space can be constructed as a strongly convergent, countable sum with respect to an orthonormal basis. By showing that the Brownian motion and Ornstein–Uhlenbeck sample paths are in the $L^2([0, 1])$ space, which is a Hilbert space, we can obtain an almost sure construction for Brownian motion and the Ornstein–Uhlenbeck process.

Our construction closely follows the Lévy–Ciesielski construction of Brownian motion. However, Lévy and Zbigniew Ciesielski use the Schauder orthonormal basis. We use the Haar orthonormal basis which is the simplest orthonormal basis in $L^2([0, 1])$. We also provide the construction for the Ornstein–Uhlenbeck process.

We will begin section 1 by providing background on measure theoretic probability, stochastic processes, Hilbert space theory and convergence. Section 2 will provide the theoretical framework for L^2 spaces and convergence in the space with respect to the Haar orthonormal basis. Section 3 will provide the almost sure construction for Brownian motion. Section 4 will provide the almost sure construction for the Ornstein–Uhlenbeck process.

1 Preliminary background

Before we begin our discussion, we will provide some necessary prerequisite definitions and theorems regarding measure theoretic probability, stochastic processes, Hilbert space theory and convergence.

1.1 Measure theoretic probability and stochastic processes.

We will begin our introduction of stochastic processes with a brief review of measure theory and probability theory.

Definition 1.1.1. In probability theory, the *sample space* Ω is an arbitrary space or set of elements ω that represent all possible outcomes on an experiment or an event.

Definition 1.1.2. A *field* \mathcal{F} on Ω (not necessarily a sample space, Ω can be any nonempty set) is a collection of subsets of Ω that satisfy the following conditions:

1. If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
2. If $A, B \in \mathcal{F}$, then $A \setminus B \in \mathcal{F}$
3. $\Omega \in \mathcal{F}$

Definition 1.1.3. A σ -*field* \mathcal{F} is a field that is closed under countable unions, meaning if $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Unless otherwise stated, let \mathcal{F} denote a σ -field going forward.

Example 1.1.1. The Borel σ -field on \mathbb{R} , $\mathcal{B}(\mathbb{R})$, is the smallest σ -field containing all open intervals in \mathbb{R} .

Definition 1.1.4. Consider a field \mathcal{F} on a space Ω . A *measure* is a set function $m : \mathcal{F} \mapsto [0, \infty]$ that satisfies the following conditions:

1. $m(\emptyset) = 0$
2. If $\{A_i\}$ is a disjoint sequence of sets in \mathcal{F} such that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, then:

$$m\left(\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}\right) = \sum_{i=1}^{\infty} m(A_i)$$

Example 1.1.2. For a field on \mathbb{R} , the Lebesgue measure λ on \mathbb{R} is defined as:

$$\lambda([a, b)) = b - a$$

Definition 1.1.5. A *measure space* is denoted (Ω, \mathcal{F}, m) , where Ω is a nonempty set, \mathcal{F} is a σ -field on Ω and m is a measure on Ω .

Definition 1.1.6. In a measure space (Ω, \mathcal{F}, m) , the *outer measure* m^* is a set function $m^* : \mathcal{P}(\Omega) \mapsto [0, \infty]$ that satisfies the following conditions:

1. $m^*(\emptyset) = 0$
2. For any $A, B \in \mathcal{P}(\Omega)$, if $A \subseteq B$ then $m^*(A) \leq m^*(B)$
3. For every sequence of sets $\{A_i\}$ of $\mathcal{P}(\Omega)$, $m^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} m^*(A_i)$

The outer measure can also be defined as,

$$m^*(A) = \inf \left\{ \sum_{i=1}^{\infty} m(A_i) \mid A_i \in \mathcal{F} \text{ for all } i, \text{ and } A \subseteq \bigcup_{i=1}^{\infty} A_i \right\}$$

and if there is no sequence of sets $\{A_i\}$ that cover A , then $m^*(A) = \infty$.

Example 1.1.3. The Lebesgue outer measure λ^* on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ is:

$$\lambda^*(A) = \inf \left\{ \sum_{i=1}^{\infty} \lambda(l_i) \mid l_i \text{ is a bounded open interval for all } i, \text{ and } A \subseteq \bigcup_{i=1}^{\infty} l_i \right\}$$

Below are a few useful theorems regarding Lebesgue measurable sets.

Lemma 1.1.1. *If a set E of \mathbb{R} is Lebesgue measurable with a finite outer measure, then for all $\epsilon > 0$, there exists an open set U such that $E \subseteq U$ and $\lambda^*(U \setminus E) < \epsilon$.*

Proof. [1] For an $\epsilon > 0$, from the definition of an outer measure, we can choose a sequence of bounded open intervals $\{l_i\}$, such that $E \subseteq \bigcup_{i=1}^{\infty} l_i$ and $\sum_{i=1}^{\infty} \lambda^*(l_i) < \lambda^*(E) + \epsilon$. Let $U = \bigcup_{i=1}^{\infty} l_i$, then $E \subseteq U$. Furthermore:

$$\lambda^*(U) \leq \sum_{i=1}^{\infty} \lambda^*(l_i) < \lambda^*(E) + \epsilon \quad (1)$$

Because $E \subseteq U$, $\lambda^*(U \setminus E) = \lambda^*(U) - \lambda^*(E) < \epsilon$ from (1). \square

Theorem 1.1.1. *Let E be a Lebesgue measurable set with a finite Lebesgue outer measure. Then, for all $\epsilon > 0$, there is a finite collection of disjoint open intervals $\{l_k\}_{k=1}^n$ such that if $\mathcal{O} = \bigcup_{k=1}^n l_k$, then:*

$$\lambda^*(E \setminus \mathcal{O}) + \lambda^*(\mathcal{O} \setminus E) = \lambda^*(E \Delta \mathcal{O}) < \epsilon$$

Proof. [3] From Lemma 1.1.1, there exists an open set U such that:

$$E \subseteq U \quad (2)$$

and

$$\lambda^*(U \setminus E) < \epsilon/2$$

Since E has finite outer measure, we can infer from the inequality in (2), $\lambda^*(U) < \lambda^*(E) + \epsilon/2$, that U also has a finite outer measure. We can decompose U as a

countable union of disjoint open intervals $U = \bigcup_{i=1}^{\infty} l_i$. Since each of these intervals are measurable (recall that the measure and outer measure for measurable intervals are equivalent), the inequality,

$$\sum_{i=1}^n \lambda(l_i) = \lambda^*\left(\bigcup_{i=1}^n l_i\right) \leq \lambda^*(U) < \infty$$

holds for all n . Thus, as we let n tend towards ∞ , $\sum_{i=1}^{\infty} \lambda(l_i) < \infty$. Because the monotonically increasing sum $\sum_{i=1}^{\infty} \lambda(l_i)$ converges by Corollary 1.3.1, there exists an n such that,

$$\lambda^*(U \setminus \mathcal{O}) = \sum_{i=n+1}^{\infty} \lambda(l_i) < \epsilon/2 \quad (3)$$

for $\mathcal{O} = \bigcup_{k=1}^n l_k$. Since $\mathcal{O} \subseteq U$, it follows that $\mathcal{O} \setminus E \subseteq U \setminus E$. It follows from (2):

$$\lambda^*(\mathcal{O} \setminus E) \leq \lambda^*(U \setminus E) < \epsilon/2 \quad (4)$$

Conversely, since $E \subseteq U$, it follows that $E \setminus \mathcal{O} \subseteq U \setminus \mathcal{O}$. It follows from (3):

$$\lambda^*(E \setminus \mathcal{O}) \leq \lambda^*(U \setminus \mathcal{O}) = \sum_{i=n+1}^{\infty} \lambda(l_i) < \epsilon/2 \quad (5)$$

Combining (4) and (5):

$$\lambda^*(\mathcal{O} \setminus E) + \lambda^*(E \setminus \mathcal{O}) = \lambda^*(E \Delta \mathcal{O}) < \epsilon$$

□

We are now ready to begin reviewing probability theory from a measure theoretic perspective.

Definition 1.1.7. A *probability measure* P on a sample space Ω is a function which assigns to each A in a σ -field \mathcal{F} on Ω a real number and satisfies the following conditions:

1. $0 \leq P(A) \leq 1$ for all $A \in \mathcal{F}$
2. $P(\emptyset) = 0, P(\Omega) = 1$
3. If A_1, A_2, \dots are disjoint sets in \mathcal{F} , then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Definition 1.1.8. If P is a probability measure on a σ -field \mathcal{F} on Ω , then (Ω, \mathcal{F}, P) is called a *probability space*.

Definition 1.1.9. A *random variable* $X(\omega)$ on a probability space (Ω, \mathcal{F}, P) is a function $X: (\Omega, \mathcal{F}, P) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ such that $X^{-1}((a, b)) \in \mathcal{F}$ for all open intervals $(a, b) \in \mathbb{R}$.

Definition 1.1.10. An *expected value* of a random variable X on a probability space (Ω, \mathcal{F}, P) , denoted as $\mathbb{E}[X]$, is the integral of X over Ω with respect to the probability measure:

$$\mathbb{E}[X] = \int_{\Omega} X dP$$

Definition 1.1.11. The *variance* of a random variable X on a probability space (Ω, \mathcal{F}, P) , denoted as $\mathbb{V}[X]$, is the integral of $(X - \mathbb{E}[X])^2$ over Ω with respect to the probability measure:

$$\mathbb{V}[X] = \int_{\Omega} (X - \mathbb{E}[X])^2 dP$$

Definition 1.1.12. The *covariance* of two random variables X_i, X_j on a probability space (Ω, \mathcal{F}, P) , denoted as $COV(X_i, X_j)$, is the integral of $(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])$ over Ω with respect to the probability measure,

$$COV(X_i, X_j) = \int_{\Omega} (X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j]) dP$$

which can be simplified as $COV(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$.

Definition 1.1.13. A *probability distribution* μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ associated with the random variable X is the measure defined as:

$$\mu = P \circ X^{-1}$$

Thus, for $A \in \mathcal{B}(\mathbb{R})$, $X^{-1}(A) \in \mathcal{F}$ and $\mu(A) = P(X^{-1}(A))$.

Definition 1.1.14. A *distribution function* $F(x)$ of the random variable X is defined as:

$$F(x) = \mu((-\infty, x]) = P(X^{-1}((-\infty, x]))$$

Below are a few useful lemmas regarding probability spaces.

Lemma 1.1.2. (First Borel–Cantelli). For a probability space (Ω, \mathcal{F}, P) , let $\{A_n\}$ be a sequence of sets in \mathcal{F} . If $\sum_{n=1}^{\infty} P(A_n)$ converges, then $P(\limsup A_n) = 0$.

Proof. [7] If $\sum_{n=1}^{\infty} P(A_n)$ converges, then $\lim_{m \rightarrow \infty} \sum_{k=m}^{\infty} P(A_k) = 0$. From $\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \subset \bigcup_{k=m}^{\infty} A_k$, it follows that:

$$P\left(\limsup_{n \rightarrow \infty} A_n\right) \leq P\left(\bigcup_{k=m}^{\infty} A_k\right) \leq \sum_{k=m}^{\infty} P(A_k)$$

Since $P(\limsup_{n \rightarrow \infty} A_n)$ is independent of m , as we let m approach ∞ , $\sum_{k=m}^{\infty} P(A_k)$ approaches 0. Thus, $P(\limsup A_n) = 0$. □

Note that $\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{A_n, i.o.\}$. Intuitively, this is the set of sample points where the event A_n occurs infinitely often as n tends toward ∞ .

Lemma 1.1.3. *Consider a probability space (Ω, \mathcal{F}, P) . Given $E \in \mathcal{F}$, for an $\alpha > 0$, if f is a non-negative measurable function, then:*

$$P(\{\omega \in E \mid f(\omega) \geq \alpha\}) \leq \frac{1}{\alpha} \int_E f dP$$

Proof. Define the set $E_\alpha = \{\omega \in E \mid f(\omega) \geq \alpha\}$. Since $f(\omega) \geq \alpha$ over the set E_α :

$$\begin{aligned} \int_{E_\alpha} \alpha dP &\leq \int_{E_\alpha} f dP \\ \int_{E_\alpha} \alpha dP &= \alpha P(E_\alpha) \leq \int_{E_\alpha} f dP \end{aligned}$$

Since $E_\alpha \in E$, it follows that:

$$P(E_\alpha) \leq \frac{1}{\alpha} \int_{E_\alpha} f dP \leq \frac{1}{\alpha} \int_E f dP$$

□

Now, we will extend the definition of a random variable to incorporate time.

Definition 1.1.15. A continuous *stochastic process* is a collection of random variables $\{X_t \mid t \geq 0\}$ indexed by time, on a common probability space.

However, stochastic processes are generally described by their finite-dimensional distributions.

Definition 1.1.16. The *finite-dimensional distributions* for a stochastic process $\{X_t \mid t \geq 0\}$ have the form $\mu_{t_1, t_2, \dots, t_n}(H) = \mathbb{P}[(X_{t_1}, \dots, X_{t_n}) \in H]$ for each finite set of times t_1, t_2, \dots, t_n in $[0, T]$ and $H \in \mathcal{B}(\mathbb{R}^n)$.

Intuitively, for a finite sequence of time points (t_1, t_2, \dots, t_n) , the vector $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ has a distribution over $\mathcal{B}(\mathbb{R}^n)$ (the n -th dimensional Borel σ -field on \mathbb{R}) defined by a probability distribution $\mu_{t_1, t_2, \dots, t_n}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$.

Below are a few useful theorems regarding stochastic processes.

Theorem 1.1.2. (Markov's inequality). *Let X be a non-negative random variable on a probability space (Ω, \mathcal{F}, P) . For an $\alpha > 0$:*

$$P(\{\omega \in \Omega \mid |X(\omega)| \geq \alpha\}) \leq \frac{1}{\alpha^k} \mathbb{E}[|X|^k]$$

Proof. [7] For a non-negative random variable X on the probability space (Ω, \mathcal{F}, P) , consider the random variable $|X|^k$ and $\alpha^k > 0$. From Lemma 1.1.3:

$$P(\{\omega \in \Omega \mid |X(\omega)|^k \geq \alpha^k\}) \leq \frac{1}{\alpha^k} \int_{\Omega} |X|^k dP \quad (6)$$

From, the definition of expected values, we can simplify (6):

$$\begin{aligned} P(\{\omega \in \Omega \mid |X(\omega)| \geq \alpha\}) &= P(\{\omega \in \Omega \mid |X(\omega)|^k \geq \alpha^k\}) \\ &\leq \frac{1}{\alpha^k} \int_{\Omega} |X|^k dP \leq \frac{1}{\alpha^k} \mathbb{E}[|X|^k] \end{aligned}$$

□

Theorem 1.1.3. (Etemadi's maximal inequality). *Let X_1, X_2, \dots, X_n be a sequence of independent random variables on the probability space (Ω, \mathcal{F}, P) . Let $S_n = \sum_{k=1}^n X_k$, and $S_0 = 0$. For $\alpha \geq 0$:*

$$P\left(\left\{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k| \geq 3\alpha\right\}\right) \leq 3 \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_k| \geq \alpha\})$$

Proof. [5] Define the set A_k to be the set where $|S_k| \geq 3\alpha$ and $|S_j| < 3\alpha$ for all $j < k$. Note that A_k is the set where $|S_j| < 3\alpha$, for all $j < k$. However, A_j is the set where $|S_j| \geq 3\alpha$. By induction, all A_k are disjoint.

Let $A = \bigcup_{k=1}^n A_k$. Consider the complement, $(\bigcup_{k=1}^n A_k)^C$. The set where $|S_k| < 3\alpha$ for all k cannot belong in any A_k , thus it is in the complement. Furthermore, consider the set where $|S_k| \geq 3\alpha$ for at least one or more k and let such k where $|S_k| \geq 3\alpha$ be denoted $k_1 < \dots < k_m$. This set necessarily belongs in A_{k_1} , since $|S_{k_1}| \geq 3\alpha$ and $|S_j| < 3\alpha$ for all $j < k_1$. Thus, $A = \{\omega \in \Omega \mid \max_{1 \leq k \leq n} |S_k| \geq 3\alpha\}$.

Given any $\omega \in A_k$, $S_k \geq 3\alpha$ or $S_k \leq -3\alpha$. Thus, if $S_n < \alpha$ or $S_n > -\alpha$, then it is necessarily true that $|S_n - S_k| > 2\alpha$. It follows that,

$$A_k \cap \{\omega \in \Omega \mid |S_n| < \alpha\} \subseteq A_k \cap \{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\} \quad (7)$$

for all $a \leq k \leq n$. Note that A_k and $\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}$ are independent. It follows that:

$$\begin{aligned} P(A) &= P(A \cap \{\omega \in \Omega \mid |S_n| \geq \alpha\}) + P(A \cap \{\omega \in \Omega \mid |S_n| < \alpha\}) \\ &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + P(A \cap \{\omega \in \Omega \mid |S_n| < \alpha\}) \end{aligned} \quad (8)$$

Since $\bigcup_{k=1}^n (A_k \cap \{\omega \in \Omega \mid |S_n| < \alpha\}) = A \cap \{\omega \in \Omega \mid |S_n| < \alpha\}$, it follows from (7) and (8) that:

$$\begin{aligned} P(A) &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + P(A \cap \{\omega \in \Omega \mid |S_n| < \alpha\}) \\ &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \sum_{k=1}^n P(A_k \cap \{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \end{aligned} \quad (9)$$

From independence and (9):

$$\begin{aligned}
P(A) &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \sum_{k=1}^n P(A_k \cap \{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \quad (10) \\
&= P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \sum_{k=1}^n P(A_k)P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \\
&\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \sum_{k=1}^n P(A_k)
\end{aligned}$$

Since A_k are disjoint and $A = \bigcup_{k=1}^n A_k$, it follows from (10):

$$\begin{aligned}
P(A) &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \sum_{k=1}^n P(A_k) \\
& \quad (11) \\
&= P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + P(A) \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\})
\end{aligned}$$

Since $P(A) \leq 1$, it follows from (11):

$$\begin{aligned}
P(A) &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + P(A) \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \\
& \quad (12) \\
&\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\})
\end{aligned}$$

Using the same logic that resulted in (7), it follows from (12) :

$$\begin{aligned}
P(A) &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_n - S_k| > 2\alpha\}) \quad (13) \\
&\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \max_{1 \leq k \leq n} (P(\{\omega \in \Omega \mid |S_n| > 2\alpha\}) \\
&\quad + P(\{\omega \in \Omega \mid |S_k| > 2\alpha\}))
\end{aligned}$$

Note that $P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) \leq \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_k| \geq \alpha\})$. Thus, it follows from (13):

$$\begin{aligned}
P(A) &\leq P(\{\omega \in \Omega \mid |S_n| \geq \alpha\}) + \max_{1 \leq k \leq n} (P(\{\omega \in \Omega \mid |S_n| > 2\alpha\}) \\
&\quad + P(\{\omega \in \Omega \mid |S_k| > 2\alpha\})) \\
&\leq 3 \max_{1 \leq k \leq n} P(\{\omega \in \Omega \mid |S_k| \geq \alpha\})
\end{aligned}$$

□

1.2 Hilbert space theory.

Definition 1.2.1. A *vector space* \mathcal{V} on a field \mathcal{F} is a set of elements that are closed under two operations, vector addition and scalar multiplication. For all $u, v, w \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$, the operations must satisfy eight specified axioms:

1. $u + v = v + u$
2. $u + (v + w) = (u + v) + w$
3. There exists a $0 \in \mathcal{V}$ such that for all $v \in \mathcal{V}$, $0 + v = v$
4. For any $v \in \mathcal{V}$, there exists a $-v \in \mathcal{V}$ such that $v + (-v) = 0$
5. $\alpha(u + v) = \alpha u + \alpha v$
6. $(\alpha + \beta)v = \alpha v + \beta v$
7. $\alpha(\beta v) = (\alpha\beta)v$
8. $1v = v$

Example 1.2.1. $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$, equipped with usual addition and scalar multiplication is a vector space.

Definition 1.2.2. Given a vector space \mathcal{V} , a *norm* on \mathcal{V} is a real-valued function $\|\cdot\| : \mathcal{V} \mapsto \mathbb{R}^+$ which satisfies the following three properties:

1. $\|x\| \geq 0$ for each $x \in \mathcal{V}$, and $\|x\| = 0$ if and only if $x = 0$
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$ for all $x \in \mathcal{V}$ and $\alpha \in \mathbb{R}$
3. (Triangle Inequality). $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathcal{V}$

Definition 1.2.3. A *normed vector space* is a vector space equipped with a norm and the metric induced by the norm, meaning that $d(x, y) = \|x - y\|$.

Definition 1.2.4. A normed vector space is *complete* if all of its Cauchy sequences converge to a point in the vector space with respect to the norm.

Definition 1.2.5. A *Banach space* \mathcal{X} is a complete normed vector space.

Example 1.2.2. $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}$ with the Euclidian norm $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ is a Banach space.

Definition 1.2.6. A real *inner product* $\langle \cdot, \cdot \rangle$ on a vector space \mathcal{V} is a function $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{R}$ satisfying the following properties:

1. (Linearity). $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$, for all $x, y, z \in \mathcal{V}$ and for all $\alpha, \beta \in \mathbb{R}$
2. $\langle x, y \rangle = \langle y, x \rangle$, for all $x, y \in \mathcal{V}$
3. For all $x \in \mathcal{V}$, $\langle x, x \rangle \geq 0$

Note, for the purposes of this paper, we will only consider real-valued inner products.

Definition 1.2.7. An *inner product space* \mathcal{X} is a normed vector space equipped with an inner product $\langle \cdot, \cdot \rangle$ and the norm is defined as:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Example 1.2.3. $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}$ with the inner product defined as the dot product ($\langle x, y \rangle = x \cdot y$) is a Hilbert space.

Definition 1.2.8. A *Hilbert space* \mathcal{H} is an inner product space that is complete with respect to the metric induced by the inner product.

Below are a few useful theorems regarding inner product spaces.

Theorem 1.2.1. (Cauchy–Schwarz inequality). For any x and y in a real inner product space:

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

Proof. From the properties of inner products:

$$0 \leq \langle \alpha x + y, \alpha x + y \rangle = \alpha^2 \langle x, x \rangle + 2\alpha \langle x, y \rangle + \langle y, y \rangle$$

Note that $\alpha^2 \langle x, x \rangle + 2\alpha \langle x, y \rangle + \langle y, y \rangle$ is a quadratic equation with respect to α that is greater than or equal to 0. Thus, it can have at most one real root and from the quadratic equation,

$$(2\langle x, y \rangle)^2 - 4\langle x, x \rangle \langle y, y \rangle \leq 0$$

it follows that:

$$\begin{aligned} \langle x, y \rangle^2 &\leq \langle x, x \rangle \langle y, y \rangle \\ |\langle x, y \rangle| &\leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} \\ |\langle x, y \rangle| &\leq \|x\| \|y\| \end{aligned}$$

□

Theorem 1.2.2. The inner product is a continuous function with respect to the induced norm in an inner product space. Meaning, for all $\epsilon > 0$, there exists $\delta_x, \delta_y > 0$ such that:

$$\|x_n - x\| < \delta_x \text{ and } \|y_n - y\| < \delta_y \text{ implies } |\langle x_n, y_n \rangle - \langle x, y \rangle| < \epsilon$$

Proof. [1] Note that the sequence $\{\|x_n\|\}$ is bounded, thus let $\|x_n\| \leq \alpha$ hold for all n . From the linearity of inner products and Cauchy–Schwarz inequality:

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n, y_n \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle| \\ &= |\langle x_n, y_n \rangle - \langle x_n, y \rangle| + |\langle x_n, y \rangle - \langle x, y \rangle| \\ &\leq |\langle x_n, y_n - y \rangle| + |\langle y, x_n - x \rangle| \\ &\leq \|x_n\| \|y_n - y\| + \|y\| \|x_n - x\| \end{aligned}$$

Thus, we can choose δ_x and δ_y such that:

$$|\langle x_n, y_n \rangle - \langle x, y \rangle| \leq \|x_n\| \|y_n - y\| + \|y\| \|x_n - x\| < \alpha \delta_y + \|y\| \delta_x = \epsilon$$

Specifically, we can choose $\delta_y = \frac{\epsilon}{2\alpha}$ and $\delta_x = \frac{\epsilon}{2\|y\|}$. □

Definition 1.2.9. Given x and y in an inner product space \mathcal{V} , x and y are *orthogonal* (denoted $x \perp y$) if $\langle x, y \rangle = 0$.

Definition 1.2.10. An *orthogonal set* is a set of vectors \mathcal{S} such that any two distinct vectors in \mathcal{S} are orthogonal. Furthermore, \mathcal{S} is called an *orthonormal set* if $\|s\| = 1$ for all $s \in \mathcal{S}$.

Definition 1.2.11. An orthogonal set \mathcal{S} is a *complete orthogonal set* in an inner product space \mathcal{X} if $x \perp s$ for all $s \in \mathcal{S}$, then $x = 0$.

We will prove the existence of a complete orthogonal set in the Hilbert space by showing that all inner product spaces have complete orthogonal sets.

Theorem 1.2.3. *Every inner product space has a complete orthogonal set*

Proof. For an inner product space, consider the set of all orthogonal sets X ordered by inclusion. X is a partially ordered set and given any chain, Y , in X , the upper bound of Y is the union of all its elements. It follows from Zorn's Lemma (see Appendix A) that X has a maximal element, $\Phi \in X$. Suppose there is a non-trivial vector v that is orthogonal to all $\phi \in \Phi$. Then $\Phi \cup \{v\}$ is an orthogonal set and $\Phi \subsetneq \Phi \cup \{v\}$. However, this cannot be the case because Φ is a maximal element. Therefore, the only vector that is orthogonal to all $\phi \in \Phi$ is the zero vector and the orthogonal set is complete. \square

Note that since all orthogonal sets can be normalized into orthonormal sets, Theorem 1.2.3 implies that all Hilbert spaces have complete orthonormal sets.

Definition 1.2.12. A complete orthonormal set in a Hilbert space is called an *orthonormal basis*.

An orthonormal basis in a Hilbert space has very nice characteristics to work with, which we will now describe.

Theorem 1.2.4. (Bessel's inequality). *If $\{\phi_i\}_{i \in I}$ is an orthonormal basis in an inner product space X , then for all $x \in X$:*

$$\sum_{i \in I} |\langle x, \phi_i \rangle|^2 \leq \|x\|^2$$

holds for all vectors x . For all x , all but an at most countable number of the $\langle x, \phi_i \rangle$ vanish.

Proof. [1] [4] Let x be a vector in an inner product space X and let $\phi_1, \phi_2, \dots, \phi_n$ be a finite orthonormal set:

$$\begin{aligned} 0 &\leq \left\langle x - \sum_{i=1}^n \langle x, \phi_i \rangle \phi_i, x - \sum_{j=1}^n \langle x, \phi_j \rangle \phi_j \right\rangle \\ &= \|x\|^2 - \sum_{i=1}^n \langle x, \phi_i \rangle \langle \phi_i, x \rangle - \sum_{j=1}^n \langle x, \phi_j \rangle \langle \phi_j, x \rangle + \sum_{i=1}^n \sum_{j=1}^n \langle \langle x, \phi_i \rangle \phi_i, \langle x, \phi_j \rangle \phi_j \rangle \\ &= \|x\|^2 - \sum_{i=1}^n |\langle x, \phi_i \rangle|^2 - \sum_{j=1}^n |\langle x, \phi_j \rangle|^2 + \sum_{i=1}^n |\langle x, \phi_i \rangle|^2 = \|x\|^2 - \sum_{i=1}^n |\langle x, \phi_i \rangle|^2 \end{aligned}$$

Thus, $\sum_{i=1}^n |\langle x, \phi_i \rangle|^2 \leq \|x\|^2$ holds for all n . Thus, $\sum_{i=1}^{\infty} |\langle x, \phi_i \rangle|^2 \leq \|x\|^2$ holds for a countable orthonormal basis.

To extend the proof to an uncountable orthonormal basis (where I is uncountable), consider the definition of an uncountable sum of positive numbers:

$$\sum_{i \in I} |\langle x, \phi_i \rangle|^2 = \sup \left\{ \sum_{i \in J} |\langle x, \phi_i \rangle|^2 \mid J \text{ is a finite subset of } I \right\}$$

From the proof above, $\|x\|^2$ is an upper bound for all finite sums $\sum_{i \in J} |\langle x, \phi_i \rangle|^2$. If $\sum_{i \in I} |\langle x, \phi_i \rangle|^2$ is the supremum (least upper bound) of all finite sums, then it follows that $\sum_{i \in I} |\langle x, \phi_i \rangle|^2 \leq \|x\|^2$. Consider the set:

$$I_n = \{i \in I \mid |\langle x, \phi_i \rangle|^2 \geq 1/n\}$$

For all $n \in \mathbb{N}$, I_n must be finite otherwise $\sum_{i \in I} |\langle x, \phi_i \rangle|^2$ would be infinite. Note that,

$$\bigcup_{i=1}^{\infty} I_n = \{i \in I \mid |\langle x, \phi_i \rangle|^2 > 0\}$$

and a countable union of finite sets is countable. This implies that at most a countable number of $\langle x, \phi_i \rangle \neq 0$. \square

Theorem 1.2.5. *An orthonormal basis $\{\phi_i\}_{i \in I}$ in a Hilbert space \mathcal{H} has the following properties:*

1. For all $x \in \mathcal{H}$, $x = \sum_{i \in I} \langle x, \phi_i \rangle \phi_i$, which converges strongly
2. (Parseval's equality). For all $x \in \mathcal{H}$, $\|x\|^2 = \langle x, x \rangle = \sum_{i \in I} |\langle x, \phi_i \rangle|^2$
3. (Generalization of Parseval's equality). For all $x, y \in \mathcal{H}$, $\langle x, y \rangle = \sum_{i \in I} \langle x, \phi_i \rangle \langle y, \phi_i \rangle$

Proof. [1] (1) Choose an arbitrary vector x in the Hilbert space. By Bessel's inequality, we know that there are at most countably many orthonormal vectors in $\{\phi_i\}_{i \in I}$ such that $\langle x, \phi_i \rangle \neq 0$. Let $i_1, i_2 \dots$ be the index for such orthonormal vectors. Also following Bessel's inequality:

$$\sum_{k=1}^{\infty} |\langle x, \phi_{i_k} \rangle|^2 \leq \|x\|^2 < \infty \quad (14)$$

Suppose $m \geq n$, by the linearity of inner products:

$$\begin{aligned} \left\| \sum_{k=n}^m \langle x, \phi_{i_k} \rangle \phi_{i_k} \right\|^2 &= \left\langle \sum_{k=n}^m \langle x, \phi_{i_k} \rangle \phi_{i_k}, \sum_{j=n}^m \langle x, \phi_{i_j} \rangle \phi_{i_j} \right\rangle \\ &= \sum_{k=n}^m \sum_{j=n}^m \langle \langle x, \phi_{i_k} \rangle \phi_{i_k}, \langle x, \phi_{i_j} \rangle \phi_{i_j} \rangle = \sum_{k=n}^m \langle x, \phi_{i_k} \rangle^2 \langle \phi_{i_k}, \phi_{i_k} \rangle \\ &+ \sum_{j \neq k} \langle x, \phi_{i_k} \rangle \langle x, \phi_{i_j} \rangle \langle \phi_{i_k}, \phi_{i_j} \rangle = \sum_{k=n}^m |\langle x, \phi_{i_k} \rangle|^2 \end{aligned} \quad (15)$$

Simplifying (15):

$$\left\| \sum_{k=1}^m \langle x, \phi_{i_k} \rangle \phi_{i_k} - \sum_{k=1}^n \langle x, \phi_{i_k} \rangle \phi_{i_k} \right\|^2 = \left\| \sum_{k=n}^m \langle x, \phi_{i_k} \rangle \phi_{i_k} \right\|^2 = \sum_{k=n}^m |\langle x, \phi_{i_k} \rangle|^2 \quad (16)$$

From (14), $\{\sum_{k=1}^n |\langle x, \phi_{i_k} \rangle|^2\}_n$ is a monotonically increasing sequence bounded above and therefore converges by Corollary 1.3.1 and is Cauchy. It follows that for all $\epsilon > 0$ there exists an N such that for all $n, m > N$:

$$\sum_{k=n}^m |\langle x, \phi_{i_k} \rangle|^2 = \left| \sum_{k=1}^m |\langle x, \phi_{i_k} \rangle|^2 - \sum_{k=1}^n |\langle x, \phi_{i_k} \rangle|^2 \right| < \epsilon \quad (17)$$

From (16) and (17) we see that the sequence $\{\sum_{k=1}^n \langle x, \phi_{i_k} \rangle \phi_{i_k}\}_n$ is Cauchy with respect to the norm and converges by completeness. Let $y = x - \sum_{k=1}^{\infty} \langle x, \phi_{i_k} \rangle \phi_{i_k}$. From the linearity of inner products,

$$\begin{aligned} \left\langle x - \sum_{k=1}^{n+m} \langle x, \phi_{i_k} \rangle \phi_{i_k}, \phi_{i_n} \right\rangle &= \langle x, \phi_{i_n} \rangle - \left\langle \sum_{k=1}^{n+m} \langle x, \phi_{i_k} \rangle \phi_{i_k}, \phi_{i_n} \right\rangle \\ &= \langle x, \phi_{i_n} \rangle - \sum_{k \neq n} \langle x, \phi_{i_k} \rangle \langle \phi_{i_k}, \phi_{i_n} \rangle - \langle x, \phi_{i_n} \rangle \langle \phi_{i_n}, \phi_{i_n} \rangle = 0 \end{aligned}$$

holds for all $m \geq 0$.

Thus, $\langle x - \sum_{k=1}^{n+m} \langle x, \phi_{i_k} \rangle \phi_{i_k}, \phi_{i_n} \rangle = 0$ for all m . Note that $\lim_{m \rightarrow \infty} \|x - \sum_{k=1}^{n+m} \langle x, \phi_{i_k} \rangle \phi_{i_k} - y\| = 0$. Thus, from the continuity of inner products, $\langle y, \phi_{i_n} \rangle = 0$ for all n . Thus, $\langle y, \phi_i \rangle = 0$ for all $i \in I$. By the definition of an orthonormal basis, $y = 0$ and thus $x = \sum_{i \in I} \langle x, \phi_i \rangle \phi_i$.

(3) Let $x = \sum_{i \in I} \langle x, \phi_i \rangle \phi_i$ and $y = \sum_{j \in I} \langle y, \phi_j \rangle \phi_j$, then by the joint continuity of the inner product:

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_{i \in I} \langle x, \phi_i \rangle \phi_i, \sum_{j \in I} \langle y, \phi_j \rangle \phi_j \right\rangle \\ &= \sum_{i \in I} \sum_{j \in I} \langle x, \phi_i \rangle \langle y, \phi_j \rangle \langle \phi_i, \phi_j \rangle \\ &= \sum_{i \in I} \langle x, \phi_i \rangle \langle y, \phi_i \rangle \end{aligned}$$

(2) Proof follows (3) but substitutes x for y . □

The significance of an orthonormal basis in a Hilbert space is that we can now construct any vector in the Hilbert space as a strongly convergent sum from Theorem 1.2.5(1).

1.3 Convergence.

Since we are working closely with convergence, this section will present a few key definitions and theorems.

Definition 1.3.1. A sequence of vectors $\{v_i\}$ in a normed vector space \mathcal{V} *converges strongly* to v if for all $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that $\|v_n - v\| < \epsilon$ for all $n > N$.

Strong convergence can be denoted as $\lim v_n = v$, $\lim v_n \rightarrow v$, or $v_n \uparrow v$ (for monotonic series).

Example 1.3.1. The sequence of real numbers $\{\frac{1}{n}\}_{n=1}^{\infty}$ converges to 0 with respect to the norm induced by the absolute value.

Definition 1.3.2. A sequence of vectors $\{v_i\}$ in a normed vector space \mathcal{V} is *Cauchy* if for all $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that $\|v_n - v_m\| < \epsilon$ for all $n, m > N$.

Now we consider the convergence of random variables.

Definition 1.3.3. A sequence of random variables $\{X_n\}$ on a probability space (Ω, \mathcal{F}, P) is said to *converge weakly* or *converge in distribution* to X if for all x points of continuity on the distribution functions $F_n(x), F(x)$, for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n > N$:

$$|F_n(x) - F(x)| < \epsilon$$

Definition 1.3.4. A sequence of random variables $\{X_n\}$ on a probability space (Ω, \mathcal{F}, P) *converges in probability* if for all $\delta > 0$, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n > N$:

$$|P(\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| \geq \delta\}) - 0| < \epsilon$$

Definition 1.3.5. A sequence of random variables $\{X_n\}$ on a probability space (Ω, \mathcal{F}, P) *converges almost surely* (denoted a.s.) if there exists a subset $S \subseteq \Omega$ where $P(S) = 1$, such that for all $\omega \in S$, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n > N$:

$$|X_n(\omega) - X(\omega)| < \epsilon$$

The most desired type of convergence for random variables is almost sure convergence. However, while random variables have values in the space of real numbers, stochastic processes have values in a space of functions. Thus, our almost sure construction will provide convergent limits in a vector space of functions for each sample point. Thus, convergence of stochastic processes connects almost sure convergence of random variables and strong convergence of vectors.

Below are a few useful convergence theorems.

Theorem 1.3.1. (Levi's). For a sequence of Lebesgue integrable functions $\{f_n\}$ that is monotonically increasing almost everywhere (denoted a.e. meaning everywhere except on a set of measure 0) and $\lim_{n \rightarrow \infty} \int f_n d\lambda < \infty$, there exists an integrable function f such that $f_n \uparrow f$ a.e. and $\int f_n d\lambda \uparrow \int f d\lambda$.

Proof. [1] Without loss of generality, we can replace $\{f_n\}$ with $\{f_n - f_1\}$ so that $f_n \geq 0$ a.e. holds for all n and $\{f_n\}$ is monotonically increasing. Let's denote $I = \lim_{n \rightarrow \infty} \int f_n d\lambda$. For all $x \in X$, define a function g such that $g(x) = \lim_{n \rightarrow \infty} f_n(x)$. Define the set:

$$E = \{x \in X \mid g(x) = \infty\} = \bigcap_{i=1}^{\infty} \bigcup_{n=1}^{\infty} \{x \in X \mid f_n(x) > i\}$$

Since f_n is a measurable function, $\{x \in X \mid f_n(x) > i\}$ is measurable by definition, and thus E is measurable. Our goal is to show that $\lambda^*(E) = 0$.

Since f_n is non-negative and measurable, there exists a sequence of simple functions $\{s_n^i\}$ such that $0 \leq s_n^i \uparrow f_n$ a.e. for all i . Let \vee denote the maximum. For each n , let $S_n = \vee_{i=1}^n s_n^i$. Notice that $s_n^i \uparrow f_n \uparrow g$. If we let both n and i to tend towards ∞ , s_n^i converges to g . Thus, $S_n \uparrow g$ a.e. and $\lim_{n \rightarrow \infty} \int S_n d\lambda = \lim_{n \rightarrow \infty} \int f_n d\lambda = I$.

Let \wedge denote the minimum. For all k , the sequence of step functions $\{S_n \wedge k \mathbb{1}_E\}$ satisfies $S_n \wedge k \mathbb{1}_E \uparrow k \mathbb{1}_E$ a.e. It follows that $\int k \mathbb{1}_E d\lambda = k \lambda^*(E) \leq \lim_{n \rightarrow \infty} \int S_n d\lambda = I < \infty$. It follows that $\lambda^*(E) = 0$.

Define a function f :

$$f(x) = \begin{cases} g(x) & \text{if } x \notin E \\ 0 & \text{if } x \in E \end{cases}$$

Thus, $f_n \uparrow f$ a.e.

From above, $S_n \leq f_n$ a.e. for all n , it follows that:

$$\lim_{n \rightarrow \infty} \int S_n d\lambda \leq \lim_{n \rightarrow \infty} \int f_n d\lambda \quad (18)$$

Fixing i , $s_n^i \leq S_n$ for all $n \geq i$. Therefore, $\int f_i d\lambda = \lim_{n \rightarrow \infty} \int s_n^i d\lambda \leq \lim_{n \rightarrow \infty} \int S_n d\lambda$ for all i . As we allow i to approach ∞ :

$$\lim_{n \rightarrow \infty} \int f_n d\lambda \leq \lim_{n \rightarrow \infty} \int S_n d\lambda \quad (19)$$

From (18) and (19), we have:

$$\lim_{n \rightarrow \infty} \int f_n d\lambda = \lim_{n \rightarrow \infty} \int S_n d\lambda = \int f d\lambda$$

□

Corollary 1.3.1. If $\{a_n\}$ is a bounded sequence of monotonically increasing real numbers then $\{a_n\}$ converges.

Proof. Let α be the upper bound of $\{a_n\}$. Define a sequence of functions $\{f_{a_n}(x)\}$ such that $f_{a_n}(x) = a_n$ for all x . Note that,

$$\int f_{a_n} d\lambda = \frac{(a_n)^2}{2} + C \leq \frac{\alpha^2}{2} + C < \infty \quad (20)$$

holds for all n since both α and C are finite. It follows from (20) that $\lim_{n \rightarrow \infty} \int f_{a_n} d\lambda < \infty$. Thus, from Theorem 1.3.1 there exists an $f(x)$ such that $f_{a_n}(x) \uparrow f(x)$. Fix an x , then $\{f_{a_n}(x)\}$ is equivalent to $\{a_n\}$ and thus $\{a_n\}$ converges to $f(x)$. \square

Theorem 1.3.2. (Lebesgue Dominated Convergence). *In the Lebesgue measure space, fix an integrable function g . If $\{f_n\}$ is a sequence of integrable functions satisfying $|f_n| \leq g$ a.e. for all n and $f_n \rightarrow f$ a.e., then f is an integrable function and:*

$$\lim_{n \rightarrow \infty} \int f_n d\lambda = \int \lim_{n \rightarrow \infty} f_n d\lambda = \int f d\lambda$$

Proof. [1] Since $|f_n| \leq g$ a.e. and $f_n \rightarrow f$ a.e., it follows that $|f| \leq g$ a.e. Thus, f is integrable because it is bounded above and below by integrable functions. Consider the sequence of functions $\{g - f_n\}$ which is a sequence of integrable functions that are greater than 0. Note that $\liminf(g - f_n) = g - f$ a.e. Thus, it follows from Fatou's Lemma (see Appendix B) that:

$$\begin{aligned} \int g d\lambda - \int f d\lambda &= \int (g - f) d\lambda = \int \liminf (g - f_n) d\lambda \\ &\leq \liminf \int (g - f_n) d\lambda = \int g d\lambda - \limsup \int f_n d\lambda \end{aligned}$$

So:

$$\int f d\lambda \geq \limsup \int f_n d\lambda \quad (21)$$

Similarly, the sequence $\{g + f_n\}$ yields:

$$\int f d\lambda \leq \liminf \int f_n d\lambda \quad (22)$$

Combining (21) and (22) yields $\limsup \int f_n d\lambda \leq \int f d\lambda \leq \liminf \int f_n d\lambda$. However, $\limsup \int f_n d\lambda \geq \liminf \int f_n d\lambda$. Thus, $\lim_{n \rightarrow \infty} \int f_n d\lambda = \int f d\lambda$. \square

2 Strong convergence of functions in the $L^2([0, 1])$ space via Haar basis

A special space that we are interested in is the $L^2([0, 1])$ space with respect to the Lebesgue measure. We will introduce this space by first defining broader the L^p space with respect to the Lebesgue measure.

2.1 L^p space.

Definition 2.1.1. In the Lebesgue measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, for real-valued functions, define the equivalence relation $f \sim g$ if $f = g$ a.e. For $1 \leq p < \infty$, let the $L^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ space (abbreviated L^p) be the collection of such equivalence classes of measurable real-valued functions f for which $|f|^p$ is integrable.

The L^p space is a normed vector space with the norm:

$$\|f\|_p = \left(\int |f|^p d\lambda \right)^{1/p}$$

We now want to show completeness of L^p with respect to $\|\cdot\|_p$.

Theorem 2.1.1. (Reisz-Fischer). *The L^p space is complete.*

Proof. [1] Let $\{f_i\}$ be any arbitrary Cauchy sequence in L^p . Thus, for any $\epsilon > 0$, there exists N such that for all $l, m > N$:

$$\|f_l - f_m\|_p < \epsilon$$

By passing to a subsequence, we know that for all n , we can choose $i_{n+1}, i_n > N_n$ such that:

$$\|f_{i_{n+1}} - f_{i_n}\|_p < 2^{-n}$$

Note that $\sum_{n=1}^{\infty} \|f_{i_{n+1}} - f_{i_n}\|_p$ converges to a value less than or equal to 1. It suffices to prove that if the subsequence f_{i_n} is norm convergent, then f_i is norm convergent as well.

Construct a sequence of monotonically increasing functions $\{g_n\}$ in L^p such that $g_1 = 0$ and for $n \geq 2$:

$$g_n = |f_{i_1}| + \sum_{j=2}^n |f_{i_j} - f_{i_{j-1}}|$$

By the triangle inequality of norms,

$$\begin{aligned} \int |f_{i_n}|^p d\lambda &\leq \int (g_n)^p d\lambda = (\|g_n\|_p)^p \\ &\leq \left(\|f_{i_1}\|_p + \sum_{j=2}^{\infty} \|f_{i_j} - f_{i_{j-1}}\|_p \right)^p \\ &\leq \left(\|f_{i_1}\|_p + 1 \right)^p \end{aligned}$$

holds for all n and p . In the case of $p = 1$, we can see that $\int g_n d\lambda$ is bounded above for all n . From Levi's Theorem, there exists some $g \in L^p$ such that $g_n \uparrow g$ a.e.

From the triangle inequality, assuming $m \geq n$:

$$|f_{i_m} - f_{i_n}| = \left| \sum_{j=n+1}^m (f_{i_j} - f_{i_{j-1}}) \right| \leq \sum_{j=n+1}^m |f_{i_j} - f_{i_{j-1}}| = g_m - g_n = |g_m - g_n| \quad (23)$$

Since $\{g_n\}$ converges in the reals, it must be Cauchy as well. It follows from (23) that for any $\epsilon > 0$, there exists $m, n > N$ such that:

$$|f_{i_m} - f_{i_n}| \leq |g_m - g_n| < \epsilon$$

Thus, $\{f_{i_n}\}$ is Cauchy in the reals as well. Since all Cauchy sequences in the reals converge, $\{f_{i_n}\}$ converges pointwise a.e. to an f . Since the sequence of functions $\{|f_{i_n}|^p\}$ are integrable and bounded above by an integrable g , it follows from the Lebesgue Dominated convergence theorem that $|f|^p$ is integrable and $f \in L^p$ by definition.

Since $|f_{i_n}| \leq g_n \leq g$ for all n , $|f - f_{i_n}| \leq |f| + |f_{i_n}| \leq 2g$ a.e. and thus $|f - f_{i_n}|^p$ is bounded above by an integrable function $|2g|^p$. Furthermore, since f_{i_n} converges pointwise a.e. to f , we can say that $\lim |f_{i_n} - f| = 0$ a.e., and it follows that $\lim |f_n - f|^p = 0$ a.e. By the Lebesgue Dominated Convergence Theorem:

$$\begin{aligned} \lim \|f - f_{i_n}\|_p &= \left(\lim \int |f - f_{i_n}|^p d\lambda \right)^{1/p} = \left(\int \lim |f - f_{i_n}|^p d\lambda \right)^{1/p} \\ &= \left(\int 0 d\lambda \right)^{1/p} = 0 \end{aligned}$$

We conclude that $\lim_{n \rightarrow \infty} \|f - f_{i_n}\|_p = 0$. □

2.2 L^2 space.

Consider a special instance of the L^p space where $p = 2$, the L^2 space. The L^2 space has a norm $\|\cdot\|_2$, which we will denote as $\|\cdot\|$ going forward. The norm of the L^2 space is defined as,

$$\|f\|^2 = \int f^2 d\lambda$$

and inner product defined as:

$$\langle f, g \rangle = \int fg d\lambda$$

Obviously, the L^2 space is an inner product space. From Theorem 2.1.1, it follows that the L^2 space is a Hilbert space. From Theorem 1.2.5(1), there exists an orthonormal basis $\{\phi_i\}_{i \in I}$ in L^2 such that:

$$f = \sum_{i \in I} \langle f, \phi_i \rangle \phi_i = \sum_{i \in I} \left[\phi_i \int f \phi_i d\lambda \right]$$

This is nice because we can construct any function $f \in L^2$ as a strongly convergent, countable sum.

2.3 Haar basis.

The goal now is to construct an orthonormal basis in L^2 . A common construction of orthonormal basis arrives from wavelet systems.

Definition 2.3.1. A wavelet system in $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ is a collection of functions $\{\phi_{j,k}\}_{j,k \in \mathbb{Z}^+ \cup \{0\}}$ such that:

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)$$

For some fixed function $\phi \in L^2$, ϕ is referred to as the *mother wavelet*. All following functions in the wavelet family are either translations (by shifting k) or dilations (by increasing j) of the mother wavelet.

A wavelet system that we are interested in is the Haar basis. The Haar basis is an orthonormal wavelet basis for $L^2([0, 1])$. For the purposes of this paper, it is fine to constrict our measure space to the interval $[0, 1]$ since the functions that we are concerned with can be defined on the interval $[0, 1]$. We will define the Haar basis and show that it is an orthonormal basis for $L^2([0, 1])$. We begin by defining dyadic intervals.

Definition 2.3.2. The *dyadic intervals* on the interval $[0, 1]$ is the collection of intervals:

$$\{I_{j,k} = [\frac{k}{2^j}, \frac{k+1}{2^j}) \mid j, k \in \mathbb{Z}^+ \cup \{0\} \text{ and } k \leq 2^j - 1\}$$

Definition 2.3.3. The *Haar basis* $h_{j,k}(x)$, for $j, k \in \mathbb{Z}^+ \cup \{0\}, k \leq 2^j - 1$, is defined as:

$$h_{j,k}(x) = \begin{cases} 2^{j/2} & \text{if } x \in [\frac{k}{2^j}, \frac{k+\frac{1}{2}}{2^j}) \\ -2^{j/2} & \text{if } x \in [\frac{k+\frac{1}{2}}{2^j}, \frac{k+1}{2^j}) \\ 0 & \text{otherwise} \end{cases}$$

Below is an illustration of Haar basis functions:

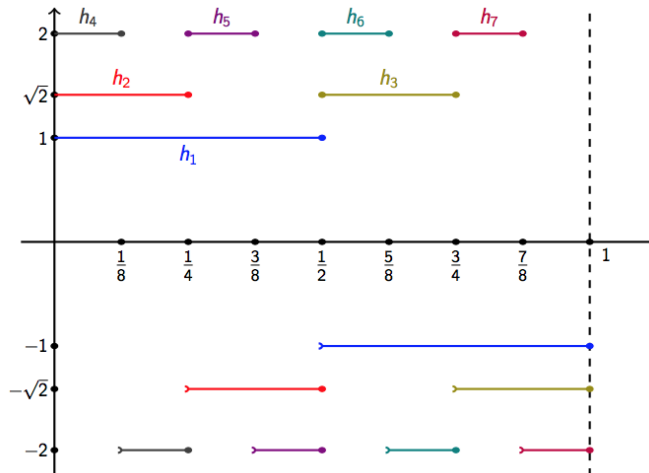


Figure 1: [8] Haar basis functions

The Haar basis has support on a dyadic interval $I_{j,k}$ determined by j and k , where j represents the dyadic level and k represents the position. The Haar realizes positive values on the first half of the support and negative values on the second half.

To continue, we must show that the Haar basis is an orthonormal basis in $L^2([0, 1])$.

Theorem 2.3.1. *The Haar basis is an orthonormal basis in $L^2([0, 1])$.*

Proof. This proof will consist of three parts:

1. Show that the Haar basis functions are normal
2. Show that the Haar basis functions are orthogonal
3. Show that the Haar basis functions are dense in $L^2([0, 1])$

Part 1. Note that,

$$(h_{j,k})^2(x) = 2^j \mathbb{1}_{[k2^{-j}, (k+\frac{1}{2})2^{-j})}(x) + (2^j) \mathbb{1}_{[(k+\frac{1}{2})2^{-j}, (k+1)2^{-j})}(x)$$

is a simple function. Since $(h_{j,k})^2$ is a non-negative simple function, it follows that:

$$\int_0^1 (h_{j,k})^2(x) d\lambda = (2^j) \lambda([\frac{k}{2^j}, \frac{k+\frac{1}{2}}{2^j})) + (2^j) \lambda([\frac{k+\frac{1}{2}}{2^j}, \frac{k+1}{2^j})) = 1$$

From the definition L^2 -norm, it follows that:

$$\|h_{j,k}\| = \sqrt{\int_0^1 (h_{j,k})^2(x) d\lambda} = 1$$

Thus, the Haar basis consists of unit vectors.

Part 2. Suppose there are two arbitrary Haar basis functions, $h_{j,k}$ and $h_{j',k'}$, such that $(j, k) \neq (j', k')$. There are two cases:

1. Assume $I_{j,k} \subset I_{j',k'}$, therefore the support of $h_{j,k}$ is contained in the support of $h_{j',k'}$. Thus, $h_{j,k}(x)h_{j',k'}(x)$ will only have support on $I_{j,k}$ since $h_{j,k}(x) = 0$ for all $x \notin I_{j,k}$. Furthermore, $h_{j',k'}$ will take on a constant value of $\pm 2^{j'/2}$ on $I_{j,k}$ since $h_{j',k'}$ exists on the coarser dyadic level $I_{j',k'}$. It follows that:

$$\langle h_{j,k}, h_{j',k'} \rangle = \pm 2^{j'/2} \int_0^1 h_{j,k}(x) d\lambda \quad (24)$$

Since $h_{j,k}$ is a simple function, we can decompose $h_{j,k}$ as $h_{j,k} = h_{j,k}^+ - h_{j,k}^-$, where:

$$\begin{aligned} h_{j,k}^+(x) &= \max\{h_{j,k}(x), 0\} \\ h_{j,k}^-(x) &= \max\{-h_{j,k}(x), 0\} \end{aligned}$$

Thus, we can integrate $h_{j,k}$ as such:

$$\begin{aligned} \int_0^1 h_{j,k}(x) d\lambda &= \int_0^1 h_{j,k}^+(x) d\lambda - \int_0^1 h_{j,k}^-(x) d\lambda \\ &= (2^{j/2})\lambda\left(\left[\frac{k}{2^j}, \frac{k+\frac{1}{2}}{2^j}\right)\right) - (2^{j/2})\lambda\left(\left[\frac{k+\frac{1}{2}}{2^j}, \frac{k+1}{2^j}\right)\right) = 0 \end{aligned} \quad (25)$$

From (24) and (25), $\langle h_{j,k}, h_{j',k'} \rangle = 0$. The case for $I_{j',k'} \subset I_{j,k}$ can be proven in a similar fashion.

2. Assume $I_{j,k} \cap I_{j',k'} = \emptyset$, therefore the support of $h_{j,k}$ and $h_{j',k'}$ are disjoint. Therefore, $h_{j,k}(x)h_{j',k'}(x) = 0$. It follows that:

$$\langle h_{j,k}, h_{j',k'} \rangle = \int_0^1 h_{j,k}(x)h_{j',k'}(x) d\lambda = \int_0^1 0 d\lambda = 0$$

Thus, distinct Haar basis functions are orthogonal.

Part 3. To show the density of Haar functions in $L^2([0,1])$, we will take three steps:

1. Show that simple functions are dense in L^2
2. Show that dyadic step functions are dense in simple functions
3. Show that Haar basis functions are dense in dyadic step functions on the interval $[0,1]$

Step 1. Consider any real-valued measurable function $f \in L^2$. Consider a sequence of simple functions $\{f_n(\omega)\}$ defined as:

$$f_n(\omega) = \begin{cases} -n & \text{if } -\infty \leq f(\omega) \leq -n \\ -(k-1)2^{-n} & \text{if } -k2^{-n} \leq f(\omega) \leq -(k-1)2^{-n}, \text{ for } 1 \leq k \leq n2^n \\ (k-1)2^{-n} & \text{if } (k-1)2^{-n} \leq f(\omega) \leq k2^{-n}, \text{ for } 1 \leq k \leq n2^n \\ n & \text{if } n \leq f(\omega) \leq \infty \end{cases}$$

Suppose $f(\omega) \geq 0$ and $(k-1)2^{-n} \leq f(\omega) < k2^{-n}$. Then, $f_n(\omega) = (k-1)2^{-n}$. Furthermore by looking at the next dyadic level, either:

$$(2k-2)2^{-(n+1)} \leq f(\omega) < (2k-1)2^{-(n+1)}$$

or

$$(2k-1)2^{-(n+1)} \leq f(\omega) < 2k2^{-(n+1)}$$

Thus, $f_{n+1}(\omega)$ is either equal to $(2k-2)2^{-(n+1)}$ or $(2k-1)2^{-(n+1)}$. Therefore, $f_n(\omega) \leq f_{n+1}(\omega)$ holds for all n . It follows that $f_n(\omega)$ is monotonically increasing when $f(\omega) \geq 0$. It can be shown in a similar fashion that $f_n(\omega)$ is monotonically decreasing when $f(\omega) \leq 0$.

Again, suppose $f(\omega) \geq 0$ and $(k-1)2^{-n} \leq f(\omega) < k2^{-n}$. Then, $f_n(\omega) = (k-1)2^{-n}$. It follows that $|f_n(\omega) - f(\omega)| \leq 2^{-n}$ holds for all n . Therefore, for all $\epsilon > 0$, there exists a N such that for all $n > N$, $|f_n(\omega) - f(\omega)| < \epsilon$. Therefore, if $f(\omega) \geq 0$ then $0 \leq f_n(\omega) \uparrow f(\omega)$. It can be shown in a similar fashion that if $f(\omega) \leq 0$ then $0 \geq f_n(\omega) \downarrow f(\omega)$.

It follows that f_n^+ converges pointwise to f^+ and f_n^- converges pointwise to f^- . Since $|f^+ - f_n^+|^2 \leq |f^+|^2$, then by the Lebesgue Dominated Convergence Theorem:

$$\lim_{n \rightarrow \infty} \int |f^+ - f_n^+|^2 d\lambda = \int \lim_{n \rightarrow \infty} |f^+ - f_n^+|^2 d\lambda = 0$$

In other words, $\lim_{n \rightarrow \infty} \|f^+ - f_n^+\| = 0$. The same can be shown for $\lim_{n \rightarrow \infty} \|f_n^- - f^-\| = 0$. It follows that:

$$\begin{aligned} \lim_{n \rightarrow \infty} \|f - f_n\| &= \lim_{n \rightarrow \infty} \|f^+ - f^- - (f_n^+ - f_n^-)\| \\ &\leq \lim_{n \rightarrow \infty} \|f^+ - f_n^+\| + \lim_{n \rightarrow \infty} \|f_n^- - f^-\| = 0 \end{aligned} \quad (26)$$

Thus, it follows from (26) that the simple functions $\{f_n\}$ are dense in L^2 .

Step 2. Since f_n is a simple function, we can represent it as,

$$f_n = \sum_{i=1}^{n2^{n+1}} a_i \mathbb{1}_{A_i}(x) + n \mathbb{1}_{[f(\omega) \geq n]} - n \mathbb{1}_{[f(\omega) \leq -n]}$$

where A_i are measurable, disjoint sets. Let $b = \vee_{i=1}^{n2^{n+1}} a_i$. From Theorem 1.1.1, for all $\epsilon > 0$, there exists a finite collection of disjoint open intervals $\{l_i\}_{i=1}^m$ such that if $\mathcal{O}_i = \bigcup_{i=1}^m l_i$, then:

$$\lambda^*(\mathcal{O}_i \triangle A_i) < \frac{\epsilon}{4b^2 n 2^{n+1}} \quad (27)$$

Let the left and right endpoints of l_i be denoted as α_i and β_i , respectively. Consider the set of all dyadic intervals $\{I_{j,k}\}_{j,k \in \mathbb{Z}^+ \cup \{0\}}$ on \mathbb{R} . Fix j and define d_i^j to be a union of dyadic intervals,

$$d_i^j = \bigcup_{t=1}^T [(k_i + (t-1))2^{-j}, (k_i + t)2^{-j}] = [k_i 2^{-j}, (k_i + T)2^{-j}]$$

such that:

$$\begin{aligned} (k_i - 1)2^{-j} &< \alpha_i \leq k_i 2^{-j} \\ \text{and} \\ (k_i + T)2^{-j} &\leq \beta_i < (k_i + T + 1)2^{-j} \end{aligned}$$

Denote $\mathcal{D} = [(k_i - 1)2^{-j}, (k_i + T + 1)2^{-j}]$. Note that $d_i^j \subseteq l_i \subseteq \mathcal{D}$ and $\lambda(d_i^j \triangle \mathcal{D}) = 2 \frac{1}{2^j}$.

It follows from the set relations that $d_i^j \Delta l_i \subseteq d_i^j \Delta \mathcal{D}$. Thus:

$$\lambda(d_i^j \Delta l_i) \leq \lambda(d_i^j \Delta \mathcal{D}) = 2 \frac{1}{2^j}$$

Since d_i^j is contained in l_i which are disjoint, d_i^j is disjoint for all i and $\bigcup_{i=1}^m (d_i^j \Delta l_i) = \mathcal{O}_i \Delta \bigcup_{i=1}^m d_i^j$. It follows from the countable additivity of measures that:

$$\lambda\left(\mathcal{O}_i \Delta \bigcup_{i=1}^m d_i^j\right) = \lambda\left(\bigcup_{i=1}^m (d_i^j \Delta l_i)\right) = \sum_{i=1}^m \lambda(d_i^j \Delta l_i) \leq m \frac{2}{2^j}$$

Given m , we can choose j such that:

$$\lambda\left(\mathcal{O}_i \Delta \bigcup_{i=1}^m d_i^j\right) < \frac{\epsilon}{4b^2 n 2^{n+1}} \quad (28)$$

Since both \mathcal{O}_i and A_i are measurable, we can treat the outer measure as the measure. By representing the measure of the symmetric differences as the L^2 -norm of the differences of indicator functions:

$$\begin{aligned} \sqrt{\lambda\left(\bigcup_{i=1}^m d_i^j \Delta A_i\right)} &= \|\mathbb{1}_{\bigcup_{i=1}^m d_i^j} - \mathbb{1}_{A_i}\| \\ &\leq \|\mathbb{1}_{\bigcup_{i=1}^m d_i^j} - \mathbb{1}_{\mathcal{O}_i}\| + \|\mathbb{1}_{\mathcal{O}_i} - \mathbb{1}_{A_i}\| \\ &= \sqrt{\lambda\left(\mathcal{O}_i \Delta \bigcup_{i=1}^m d_i^j\right)} + \sqrt{\lambda(\mathcal{O}_i \Delta A_i)} \end{aligned} \quad (29)$$

It follows from (27), (28) and (29) that

$$\lambda\left(\bigcup_{i=1}^m d_i^j \Delta A_i\right) < \frac{\epsilon}{b^2 n 2^{n+1}} \quad (30)$$

Define a dyadic step function:

$$g_n = \sum_{i=1}^{n2^{n+1}} a_i \mathbb{1}_{\bigcup_{i=1}^m d_i^j}(x) + n \mathbb{1}_{[f(\omega) \geq n]} - n \mathbb{1}_{[f(\omega) \leq -n]}$$

It follows that:

$$\begin{aligned}
\|f_n - g_n\| &= \int \left| \sum_{i=1}^{n2^{n+1}} a_i \mathbb{1}_{A_i}(x) - \sum_{i=1}^{n2^{n+1}} a_i \mathbb{1}_{\bigcup_{i=1}^m d_i^j}(x) \right|^2 d\lambda & (31) \\
&= \int \left| \sum_{i=1}^{n2^{n+1}} a_i (\mathbb{1}_{A_i}(x) - \mathbb{1}_{\bigcup_{i=1}^m d_i^j}(x)) \right|^2 d\lambda \\
&\leq \int \sum_{i=1}^{n2^{n+1}} |a_i (\mathbb{1}_{A_i}(x) - \mathbb{1}_{\bigcup_{i=1}^m d_i^j}(x))|^2 d\lambda \\
&= \int \sum_{i=1}^{n2^{n+1}} (a_i)^2 |\mathbb{1}_{A_i}(x) - \mathbb{1}_{\bigcup_{i=1}^m d_i^j}(x)| d\lambda = \sum_{i=1}^{n2^{n+1}} (a_i)^2 \lambda \left(\bigcup_{i=1}^m d_i^j \triangle A_i \right) \\
&\leq b^2 \sum_{i=1}^{n2^{n+1}} \lambda \left(\bigcup_{i=1}^m d_i^j \triangle A_i \right)
\end{aligned}$$

From (30) and (31), it follows that for all $\epsilon > 0$, there exists a g_n such that,

$$\|f_n - g_n\| \leq b^2 \sum_{i=1}^{n2^{n+1}} \lambda(\bigcup_{i=1}^m d_i^j \triangle A_i) < b^2 n 2^{n+1} \frac{\epsilon}{b^2 n 2^{n+1}} = \epsilon$$

holds for all n . As we take the limit of n to ∞ , it follows that the dyadic step functions are dense in the simple functions which are dense in L^2 . So, dyadic step functions are dense in L^2 .

Step 3. [2] Let the span of dyadic step functions at the j -th dyadic level be denoted as V_j . Thus $\lim_{j \rightarrow \infty} V_j = L^2$. Define a series of dyadic step functions $\{p_{j,k}\}$ for $j, k \in \mathbb{Z} \cup \{0\}$ and $0 \leq k \leq 2^{j+1} - 1$ on the interval $[0, 1]$ as such:

$$p_{j,k}(x) = \begin{cases} 2^{j/2} & \text{if } x \in [\frac{k}{2^j}, \frac{k+1}{2^j}) \\ 0 & \text{otherwise} \end{cases}$$

Note that $h_{j,k} = p_{j+1,2k} - p_{j+1,2k+1}$. Let W_j denote the span of the Haar basis functions at the j -th dyadic level. Consider the expansion of any arbitrary element in V_{j+1} :

$$\begin{aligned}
\sum_{k \leq 2^{j+1}-1} a_k p_{j+1,k} &= \sum_{l \leq 2^j-1} (a_{2l} p_{j+1,2l} + a_{2l+1} p_{j+1,2l+1}) \\
&= \sum_{l \leq 2^j-1} \left[\frac{a_{2l} + a_{2l+1}}{2} (p_{j+1,2l} + p_{j+1,2l+1}) \right. \\
&\quad \left. + \frac{a_{2l} - a_{2l+1}}{2} (p_{j+1,2l} - p_{j+1,2l+1}) \right] \\
&= \sum_{l \leq 2^j-1} \frac{a_{2l} + a_{2l+1}}{2} p_{j,l} + \sum_{l \leq 2^j-1} \frac{a_{2l} - a_{2l+1}}{2} h_{j,l}
\end{aligned}$$

It follows that $V_{j+1} = V_j \oplus W_j$. From induction, it follows that $V_j = \bigoplus_{i=0}^{j-1} W_i$. So, $\lim_{j \rightarrow \infty} \bigoplus_{i=0}^{j-1} W_i = L^2([0, 1])$. \square

Now that we have shown that the Haar basis is an orthonormal basis for $L^2([0, 1])$, we can use the Haar basis to construct functions in $L^2([0, 1])$.

3 Brownian motion

One stochastic process that we are interested in is Brownian motion, which is a Gaussian stochastic process. Brownian motion has traditionally been defined as a weak or distributional limit of independent random walks. However, by showing that Brownian motion sample paths are in $L^2([0, 1])$, we will be able to create an almost sure construction of a Brownian motion sample path using the Haar basis.

Definition 3.0.1. *Brownian motion* is a stochastic process $\{B_t | t \geq 0\}$, on a probability space (Ω, \mathcal{F}, P) , which satisfies the following properties:

1. $B_0 = 0$ almost surely
2. Non-overlapping increments are independent. If $0 \leq t_0 \leq t_1 \leq \dots \leq t_k$, then:

$$\begin{aligned} & P\left(\bigcap_{i \leq k} \{\omega \in \Omega \mid B_{t_i}(\omega) - B_{t_{i-1}}(\omega) \in A_i\}\right) \\ &= \prod_{i \leq k} P(\{\omega \in \Omega \mid B_{t_i}(\omega) - B_{t_{i-1}}(\omega) \in A_i\}) \quad \text{for } A_i \in \mathcal{B}(\mathbb{R}) \end{aligned}$$

3. For $0 \leq s < t$, $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$:

$$P(\{\omega \in \Omega \mid B_t(\omega) - B_s(\omega) \in A\}) = \frac{1}{\sqrt{2\pi(t-s)}} \int_A e^{\frac{-x^2}{2(t-s)}} d\lambda \quad \text{for } A \in \mathcal{B}(\mathbb{R})$$

4. The mapping $t \mapsto B_t$ is continuous almost surely (proof to follow shortly [7])

One way to also think about Brownian motion is by its finite-dimensional distributions. Let X_1, \dots, X_n be independent, Gaussian random variables with mean 0 and variances $t_1, t_2 - t_1, \dots, t_n - t_{n-1}$. Define another set of random variable S_1, \dots, S_n , where $S_i = X_1 + \dots + X_i$. Let μ_{t_i} be the probability distribution for S_i . Then, μ_{t_1, \dots, t_n} are the finite-dimensional distributions for B_n over the partition: $0 \leq t_1 \leq t_2 \leq \dots \leq t_n = t$.

Theorem 3.0.1. *The covariance of $B_s B_t$ is $t \wedge s$.*

Proof. Suppose $s < t$. Then:

$$COV(B_s, B_t) = COV(B_s, B_s - B_s + B_t) = COV(B_s, B_s) + COV(B_s, B_t - B_s)$$

Since B_s and $B_t - B_s$ are non-overlapping increments, they are independent and the last term equals 0. The first term is simply the variance of B_s , which is s . The inverse result can be shown for $t < s$. \square

3.1 Brownian motion as a sample path in $L^2([0, 1])$.

To show that the Brownian motion sample path is an element of $L^2([0, 1])$, it is sufficient to show that the Brownian motion sample path is continuous on $[0, 1]$. If B_t is continuous on a compact interval, then it is a measurable function and $\int_0^1 |B_t|^2 d\lambda < \infty$ is defined. Thus, it satisfies all the conditions of the L^2 space.

To show that the Brownian motion sample path is continuous, we will first show that it is continuous on the dyadic rationals. Then we will extend the dyadic rationals to the reals and continuity is retained.

Define D to be the set of all non-negative dyadic rationals and define $I_{n,k} = [k2^{-n}, (k+2)2^{-n})$, these are dyadic intervals of the length $\frac{2}{2^n}$. Let:

$$M_{n,k}(\omega) = \sup_{r \in I_{n,k} \cap D} |B_r(\omega) - B_{k2^{-n}}(\omega)|$$

$$M_n(\omega) = \max_{0 \leq k < n2^n} M_{n,k}(\omega)$$

Our goal is to show that $\sum_{n=0}^{\infty} P(\{\omega \in \Omega \mid M_n > n^{-1}\})$ converges. If it does, then it follows from the first Borel–Cantelli lemma that the set $A = \{\omega \in \Omega \mid M_n > n^{-1}, i.o.\}$ has probability 0. In other words, it is with probability 0 that the sequence $\{M_n\}$ takes on a value greater than n^{-1} infinitely often. Thus, it is with probability 1 that for all $t > 0$ and $\epsilon > 0$ there exists an n such that $t < n$, $2n^{-1} < \epsilon$ and $M_n(\omega) \leq n^{-1}$. Choose any two dyadic rationals, r and r' , that lie in $[0, t]$ such that $|r - r'| < \delta = 2^{-n}$. Thus, for some $k < n2^n$, r and r' are in $I_{n,k} \cap [0, t]$. It follows that:

$$\begin{aligned} |B_r(\omega) - B_{r'}(\omega)| &= |(B_r(\omega) - B_{k2^{-n}}(\omega)) + (B_{k2^{-n}}(\omega) - B_{r'}(\omega))| \\ &\leq |(B_r(\omega) - B_{k2^{-n}}(\omega))| + |(B_{k2^{-n}}(\omega) - B_{r'}(\omega))| \\ &\leq 2M_{n,k}(\omega) \leq 2M_n(\omega) \leq 2n^{-1} < \epsilon \end{aligned}$$

Therefore, B_r is uniformly continuous over the dyadic rationals in $[0, t]$, for all t .

To show that $\sum_{n=0}^{\infty} P(\{\omega \in \Omega \mid M_n > n^{-1}\})$ converges, consider the sum,

$$S_i = \sum_{j=1}^i (B_{t+\delta j 2^{-m}} - B_{t+\delta(j-1)2^{-m}}), \quad S_0 = (B_t - B_t)$$

where $\delta = \frac{2}{2^n}$, the length of the interval $I_{n,k}$. Notice that since all of the intervals are disjoint, thus the terms of S_i are all independent. Thus, from

Etemadi's maximal inequality:

$$P\left(\left\{\omega \in \Omega \mid \max_{i \leq 2^m} |S_i| > \alpha\right\}\right) \leq 3 \max_{i \leq 2^m} P(\{\omega \in \Omega \mid |S_i| \geq \alpha/3\}) \quad (32)$$

Note that S_i is equal to $(B_{t+\delta i 2^{-m}} - B_t)$. The first observation is that $\lim_{m \rightarrow \infty} \max_{i \leq 2^m} |S_i| = M_{n,t 2^n}$. As we let m tend towards ∞ , $\max_{i \leq 2^m} |S_i|$ looks at the maximum change relative to the starting position t over the dyadics of an interval of length δ . Setting $k 2^{-n} = t$, this is precisely $M_{n,t 2^n}$. The second observation is that $|S_i|$ is a Gaussian random variable with mean 0 and variance $\delta i 2^{-m}$. Thus, $|S_i|^4$ is the fourth moment of $(B_{t+\delta i 2^{-m}} - B_t)$, which is $3(\delta i 2^{-m})^2$. Thus, $\max_{i \leq 2^m} |S_i|^4 = 3\delta^2$. Thus, it follows from Markov's inequality:

$$3 \max_{i \leq 2^m} P(\{\omega \in \Omega \mid |S_i| \geq \alpha/3\}) \leq \frac{3}{(\alpha/3)^4} \max_{i \leq 2^m} \mathbb{E}[|S_i|^4] = \frac{3^6 \delta^2}{\alpha^4} \quad (33)$$

As we let m approach ∞ , from (32) and (33), we arrive at the inequality:

$$P(\{\omega \in \Omega \mid M_{n,t 2^n} > \alpha\}) \leq \frac{3^6 \delta^2}{\alpha^4}$$

Note that $\{\omega \in \Omega \mid M_n > \alpha\} \subseteq \bigcup_{k=0}^{n 2^n} \{\omega \in \Omega \mid M_{n,k} > \alpha\}$. Since if $M_n > \alpha$, then for some $0 \leq k < n 2^n$, $M_{n,k} > \alpha$. By substituting n^{-1} for α , it follows that:

$$\begin{aligned} P(\{\omega \in \Omega \mid M_n > n^{-1}\}) &\leq \sum_{\substack{t = \frac{k}{2^n} \\ 0 \leq k \leq n 2^n}} P(\{\omega \in \Omega \mid M_{n,t 2^n} > n^{-1}\}) \\ &\leq n 2^n \frac{3^6 \delta^2}{(n^{-1})^4} = \frac{4(3^6)n^5}{2^n} \end{aligned}$$

From the Cauchy ratio test, we can see that the sum $\sum_{n=0}^{\infty} P(\{\omega \in \Omega \mid M_n > n^{-1}\})$ converges.

Therefore, Brownian motion is uniformly continuous over $D \cap [0, t]$, for all t . Since the dyadic rationals are dense in the reals, it follows that Brownian motion is continuous over the reals on the interval $[0, t]$, for all t .

For the purposes of this paper, we are interested in Brownian motion on the interval $[0, 1]$, thus we let $t = 1$.

3.2 Haar basis expansion of Brownian motion.

Since the Haar basis is an orthonormal basis for $L^2([0, 1])$, let us consider the Brownian motion sample path on the interval $[0, 1]$ by fixing a $\omega \in \Omega$. Again, from the prior section, we know that the Brownian motion sample path is in $L^2([0, 1])$, for all $\omega \in \Omega$. It follows from Theorem 1.2.5(1) that for all $\omega \in \Omega$,

$$\lim_{n \rightarrow \infty} \sum_{j=0}^n \sum_{k=0}^{2^j-1} \langle B_s, h_{j,k}(s) \rangle h_{j,k}(t) = B_t$$

where the convergence is in the $L^2([0, 1])$ space. Thus, we have created an almost sure construction of the Brownian motion sample path by expanding with respect to the Haar basis.

Here, $\langle B_s, h_{j,k}(s) \rangle$ is a coefficient which we will denote as $Y_{j,k}$. From the definition of an inner product in $L^2([0, 1])$:

$$Y_{j,k} = \int_0^1 B_s h_{j,k}(s) ds$$

Since $h_{j,k}$ only has support on the dyadic interval $I_{j,k} = [k2^{-j}, (k+1)2^{-j})$, we can rewrite $Y_{j,k}$ as:

$$Y_{j,k} = 2^{j/2} \int_{\frac{k}{2^j}}^{\frac{k+\frac{1}{2}}{2^j}} B_s ds + (-2^{j/2}) \int_{\frac{k+\frac{1}{2}}{2^j}}^{\frac{k+1}{2^j}} B_s ds \quad (34)$$

Consider an evenly spaced partition of the interval $[0, t]$, $0 = t_0 < t_1 < \dots < t_k = t$, such that $t_i - t_{i-1} = \frac{1}{n}$ for some n holds for all i . If we express the integral of B_s as a Riemann sum over the partition,

$$\lim_{k \rightarrow \infty} \sum_{i=1}^k B_{t_i} (t_i - t_{i-1}) = \lim_{k \rightarrow \infty} \frac{1}{n} \sum_{i=1}^k B_{t_i}$$

we observe that the integral of B_s is really a linear combination of Gaussian random variables with mean 0, which is a Gaussian random variable centered at 0. Thus, from (34), $Y_{j,k}$ is normally distributed with mean 0. Thus, we can expand B_t as a sum of the product of Gaussian random variables and Haar basis functions:

$$B_t = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} Y_{j,k} h_{j,k}(t) \quad (35)$$

Unfortunately, the random variables $Y_{j,k}$ are dependent. They are dependent because $Y_{j,k}$ will consist of Brownian motion increments that overlap with increments that compose $Y_{j',k'}$. To get around this, we can use the Gram-Schmidt process to construct a sequence of dependent Gaussian random variables using a sequence of i.i.d. standard Gaussian random variables centered at 0. Suppose $\{X_i\}_{i=1}^{\infty}$ is a countable sequence of i.i.d. standard Gaussian random variables centered at 0. Then, we can express dependent Gaussian random variables as:

$$\begin{aligned} Y_1 &= \sigma_{1,1} X_1 \\ Y_2 &= \sigma_{2,1} X_1 + \sigma_{2,2} X_2 \\ Y_3 &= \sigma_{3,1} X_1 + \sigma_{3,2} X_2 + \sigma_{3,3} X_3 \\ &\vdots \end{aligned}$$

By determining the covariances and variances between the dependent Y_i 's, we can determine the σ coefficients. Below is an example:

Example 3.2.1.

$$\begin{aligned}\mathbb{V}[Y_1] &= \mathbb{E}[Y_1 Y_1] - \mathbb{E}[Y_1] \mathbb{E}[Y_1] \\ &= \mathbb{E}[Y_1 Y_1] = \mathbb{E}[(\sigma_{1,1})^2 (X_1)^2] \\ &= (\sigma_{1,1})^2 \mathbb{V}[X_1]\end{aligned}$$

Thus, $\sigma_{1,1} = \sqrt{\frac{\mathbb{V}[Y_1]}{\mathbb{V}[X_1]}}$. Other σ s can be calculated in a similar fashion. Thus, our goal now is to determine the covariances between the $Y_{j,k}$'s.

Let j, k, j', k' be any arbitrary $j, k, j', k' \in \mathbb{Z}^+ \cup \{0\}$ and $k \leq 2^j - 1, k' \leq 2^{j'} - 1$. Since the means of $Y_{j,k}$ and $Y_{j',k'}$ are 0, their covariance can be expressed as $\mathbb{E}[Y_{j,k} Y_{j',k'}]$, which in integral form is:

$$\mathbb{E}[Y_{j,k} Y_{j',k'}] = \mathbb{E} \left[\int_0^1 B_s h_{j',k'}(s) ds \int_0^1 B_t h_{j,k}(t) dt \right]$$

Recall that the expected value is the integral over the sample space:

$$\begin{aligned}\mathbb{E} \left[\int_0^1 B_s h_{j',k'}(s) ds \int_0^1 B_t h_{j,k}(t) dt \right] \\ = \int_{\Omega} \left(\int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) B_s B_t ds dt \right) dP\end{aligned}\tag{36}$$

From Fubini's Theorem (see Appendix C), (36) can be expressed as:

$$\begin{aligned}\int_{\Omega} \left(\int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) B_s B_t ds dt \right) dP \\ = \int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) \left(\int_{\Omega} B_s B_t dP \right) ds dt \\ = \int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) \mathbb{E}[B_s B_t] ds dt \\ = \int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) (s \wedge t) ds dt\end{aligned}\tag{37}$$

Thus, the covariance of $Y_{j,k}$ and $Y_{j',k'}$ can be expressed as (37). This makes the integration much easier. The remaining difficulty is integrating $(s \wedge t)$. To get around this, we can split the integral into:

$$\int_0^1 h_{j,k}(t) \left[\int_0^t h_{j',k'}(s) s ds + \int_t^1 h_{j',k'}(s) t ds \right] dt$$

Note that there are two distinct Haar basis functions, $h_{j,k}$ and $h_{j',k'}$. Thus, the resulting integral is dependent on the relationship between the support of the two Haar basis functions. Suppose $j \leq j'$, there are four possible cases:

1. *Disjoint case.* The Haar basis functions have disjoint support, where $\frac{k'+1}{2^{j'}} \leq \frac{k}{2^j}$ or $\frac{k+1}{2^j} \leq \frac{k'}{2^{j'}}$

2. *Positive containment case.* The Haar basis function with with the finer dyadic level is contained in the positive support of the other Haar basis function, where $j < j'$ and $\frac{k}{2^j} \leq \frac{k'}{2^{j'}} \leq \frac{k+\frac{1}{2}}{2^j} - \frac{1}{2^{j'}}$
3. *Negative containment case.* The Haar basis function with with the finer dyadic level is contained in the negative support of the other Haar basis function, where $j < j'$ and $\frac{k+\frac{1}{2}}{2^j} \leq \frac{k'}{2^{j'}} \leq \frac{k+1}{2^j} - \frac{1}{2^{j'}}$
4. *Identical support case.* The Haar basis functions have identical support, where $j = j'$ and $k = k'$

Note that the fourth case is really the variance of $Y_{j,k}$. When $j' \leq j$ the four cases can be adjusted accordingly. The covariance results for each case when $j \leq j'$ are below (the extensive calculations for each case have been omitted):

1. *Disjoint case.*

$$COV(Y_{j,k}, Y_{j',k'}) = 0 \quad (38)$$

2. *Positive containment case.*

$$COV(Y_{j,k}, Y_{j',k'}) = 2^{\frac{-6-j-5j'}{2}} (2^j - 2^{1+j'} k + 2^{1+j} k') \quad (39)$$

3. *Negative containment case.*

$$COV(Y_{j,k}, Y_{j',k'}) = 2^{\frac{-6-j-5j'}{2}} (-2^j + 2^{1+j'} + 2^{1+j'} k - 2^{1+j} k') \quad (40)$$

4. *Identical support case.*

$$COV(Y_{j,k}, Y_{j',k'}) = \mathbb{V}(Y_{j,k}) = \frac{1}{(3)2^{2(1+j)}} \quad (41)$$

With these covariances, we can now apply the Gram–Schmidt process to express each $Y_{j,k}$ as a linear combination of i.i.d. standard Gaussian random variables.

Therefore, to model a Brownian motion sample path, we can generate a sequence of i.i.d. standard Gaussian random variables to construct $Y_{j,k}$ s via the Gram–Schmidt process and approximate Brownian motion almost surely using (35).

4 Ornstein–Uhlenbeck process

Another Gaussian stochastic process that we are interested in is the Ornstein–Uhlenbeck process.

Definition 4.0.1. The *Ornstein–Uhlenbeck process* is a stochastic process $\{X_t | t \geq 0\}$, on a probability space (Ω, \mathcal{F}, P) . It is the solution to the stochastic differential equation:

$$dX_t = \theta(\mu - X_t)dt + \sigma dB_t \quad (42)$$

and satisfies the following properties:

1. The mean of X_t is μ
2. The variance of X_t is $\frac{\sigma^2}{2\theta e^{2\theta t}}(e^{2\theta t} - 1)$
3. The covariance of X_t and X_s is $\frac{\sigma^2}{2\theta e^{\theta(t+s)}}(e^{2\theta(s \wedge t)} - 1)$

Throughout the rest of this paper, we will assume $\mu = 0$.

The Ornstein–Uhlenbeck process is a mean-reverting process, where μ represents the mean. By using an integrating factor, we find that the integral form of the solution to (42) is:

$$X_t = e^{-\theta t} X_0 + \theta \mu e^{-\theta t} \int_0^t e^{\theta s} ds + \sigma e^{-\theta t} \int_0^t e^{\theta s} dB_s$$

4.1 Ornstein–Uhlenbeck process as a sample path in $L^2([0, 1])$.

To show that the Ornstein–Uhlenbeck process sample path is an element of L^2 , we want to show that each individual term of X_t is in L^2 , and it follows that their sum will also be in L^2 .

Since $e^{-\theta t} X_0 + \theta \mu e^{-\theta t} \int_0^t e^{\theta s} ds$ is continuous, it is in L^2 . Now, we must show that $\sigma e^{-\theta t} \int_0^t e^{\theta s} dB_s$ is in L^2 to prove that X_t is an element of L^2 .

It follows from the argument in Theorem 2.3.1, that simple functions are dense in L^2 . Since L^2 contains $e^{\theta s}$, we can represent $e^{\theta s}$ as a sequence of simple functions $\{f_n\}$, where $\lim_{n \rightarrow \infty} f_n = e^{\theta s}$. It follows that we can represent $\int_0^t e^{\theta s} dB_s$ as a limit of a stochastic integral of simple functions:

$$\int_0^t e^{\theta s} dB_s = \lim_{n \rightarrow \infty} \int_0^t f_n(s) dB_s$$

Let us denote $I_n = \int_0^t f_n(s) dB_s$. If we can show that the sequence $\{I_n\}$ is Cauchy in L^2 , then $\int_0^t e^{\theta s} dB_s$ is in L^2 by completeness.

$$\begin{aligned} \|I_n - I_m\|^2 &= \int_0^1 |I_n - I_m|^2 dt = \int_0^1 \left| \int_0^t (f_n(s) - f_m(s)) dB_s \right|^2 dt \quad (43) \\ &\leq \int_0^1 \left| \int_0^t |f_n(s) - f_m(s)| dB_s \right|^2 dt \end{aligned}$$

Since f_n converges, we know that the sequence is also Cauchy in the reals. Thus $|f_n(s) - f_m(s)| < \frac{\sqrt{\epsilon}}{B_t}$, for some $n, m \geq N$. Note that we are observing a Brownian motion sample path, meaning we are fixing a $\omega \in \Omega$. Thus, $B_t(\omega)$ is finite almost surely since it has a finite mean and variance. It follows from (43) that,

$$\|I_n - I_m\|^2 < \int_0^1 \left| \frac{\sqrt{\epsilon}}{B_t} \int_0^t dB_s \right|^2 dt \quad (44)$$

holds for some $n, m \geq N$.

Consider a partition of the interval $[0, t]$ such that $0 = s_0 < s_1 < s_2 < \dots < s_k = t$. We can express the stochastic integral $\int_0^t dB_s$ as:

$$\int_0^t dB_s = \lim_{k \rightarrow \infty} \sum_{i=1}^k (B_{s_i} - B_{s_{i-1}}) = \lim_{k \rightarrow \infty} (B_{s_k} - B_0)$$

Thus, as k approaches ∞ , $\int_0^t dB_s = B_t$. Following from (44):

$$\|I_n - I_m\|^2 < \int_0^1 \left| \frac{\sqrt{\epsilon}}{B_t} B_t \right|^2 dt = \int_0^1 \epsilon dt = \epsilon \quad (45)$$

Thus, $\{I_n\}$ is Cauchy with respect to the L^2 -norm and $\sigma e^{-\theta t} \int_0^t e^{\theta s} dB_s$ is in L^2 by completeness.

4.2 Haar basis expansion of the Ornstein–Uhlenbeck process.

Again, let's consider an Ornstein–Uhlenbeck sample path on the interval $[0, 1]$. From the prior section, we know that the Ornstein–Uhlenbeck sample path is in $L^2([0, 1])$.

From Theorem 1.2.5(1), we can create an almost sure construction of the Ornstein–Uhlenbeck sample path by expanding it with respect to the Haar basis:

$$X_t = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \langle X_s, h_{j,k}(s) \rangle h_{j,k}(t) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} Y_{j,k} h_{j,k}(t)$$

Again, we denote the coefficient $\langle X_s, h_{j,k}(s) \rangle$ as $Y_{j,k}$. From the definition of an inner product in $L^2([0, 1])$:

$$Y_{j,k} = \int_0^1 X_s h_{j,k}(s) ds$$

Since $h_{j,k}$ only has support on the dyadic interval $I_{j,k} = [k2^{-j}, (k+1)2^{-j}]$, we can rewrite $Y_{j,k}$ as:

$$Y_{j,k} = 2^{j/2} \int_{\frac{k}{2^j}}^{\frac{k+\frac{1}{2}}{2^j}} X_s ds + (-2^{j/2}) \int_{\frac{k+\frac{1}{2}}{2^j}}^{\frac{k+1}{2^j}} X_s ds$$

If we consider the integral of X_s as a Riemann sum, as we did for Brownian motion, then the integral of X_s is really a linear combination of Gaussian random variables. We assumed that the mean of X_s is 0. Thus, $Y_{j,k}$ is centered at 0. However, in the case that the mean of X_s is not 0, we can show using Fubini's Theorem that $Y_{j,k}$ is still centered at 0:

$$\begin{aligned}
\mathbb{E}[Y_{j,k}] &= \int_{\Omega} Y_{j,k} dP & (46) \\
&= \int_{\Omega} \left(2^{j/2} \int_{\frac{k}{2^j}}^{\frac{k+\frac{1}{2}}{2^j}} X_s ds + (-2^{j/2}) \int_{\frac{k+\frac{1}{2}}{2^j}}^{\frac{k+1}{2^j}} X_s ds \right) dP \\
&= 2^{j/2} \left(\int_{\frac{k}{2^j}}^{\frac{k+\frac{1}{2}}{2^j}} \left(\int_{\Omega} X_s dP \right) ds - \int_{\frac{k+\frac{1}{2}}{2^j}}^{\frac{k+1}{2^j}} \left(\int_{\Omega} X_s dP \right) ds \right) \\
&= 2^{j/2} \left(\int_{\frac{k}{2^j}}^{\frac{k+\frac{1}{2}}{2^j}} \mu ds - \int_{\frac{k+\frac{1}{2}}{2^j}}^{\frac{k+1}{2^j}} \mu ds \right) = 0
\end{aligned}$$

Therefore, $Y_{j,k}$ is a Gaussian random variable centered at 0. In fact, the above fact holds for any arbitrary Gaussian stochastic process for all μ .

By using the same techniques from section 3.2, the covariance of $Y_{j,k}$ and $Y_{j',k'}$ is:

$$\begin{aligned}
&\int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) \mathbb{E}[X_s X_t] ds dt & (47) \\
&= \int_0^1 \int_0^1 h_{j',k'}(s) h_{j,k}(t) \frac{\sigma^2}{2\theta e^{\theta(t+s)}} (e^{2\theta(s \wedge t)} - 1) ds dt
\end{aligned}$$

Note here that we assume $\mathbb{E}[X_t] = \mathbb{E}[X_s] = 0$, therefore we can substitute in the covariance equation of the Ornstein-Uhlenbeck process for $\mathbb{E}[X_s X_t]$ in (47). If the mean for the Ornstein-Uhlenbeck process were not 0, then the equation would change.

We can further split the integral into:

$$\begin{aligned}
&\int_0^1 h_{j,k}(t) \left[\int_0^t h_{j',k'}(s) \frac{\sigma^2}{2\theta e^{\theta(t+s)}} (e^{2\theta(s)} - 1) ds \right. \\
&\left. + \int_t^1 h_{j',k'}(s) \frac{\sigma^2}{2\theta e^{\theta(t+s)}} (e^{2\theta(t)} - 1) ds \right] dt
\end{aligned}$$

Again, there are four possible cases for the relationship between the support of the two Haar basis functions, the same as the ones described in section 3.2.

The covariance results for each case when $j \leq j'$ are below (the extensive calculations are omitted):

1. *Disjoint case.*

$$\begin{aligned}
&COV(Y_{j,k}, Y_{j',k'}) = \\
&\left(-\frac{\sigma^2}{\theta^3} \right) 2^{\frac{-2+j+j'}{2}} e^{-2^{-j-j'} \theta (2^j + 2^{j'} + 2^{j'} k + 2^j k')} (-1 + e^{2^{-1-j} \theta})^2 \\
&\times (-1 + e^{2^{-1-j'} \theta})^2 (1 + e^{2^{-j'} \theta (1+2k')})
\end{aligned}$$

2. *Positive containment case.*

$$\begin{aligned} \text{COV}(Y_{j,k}, Y_{j',k'}) &= \\ & \left(-\frac{\sigma^2}{\theta^3}\right) 2^{\frac{-2+j+j'}{2}} e^{-2^{-j-j'}\theta(2^j+2^{j'}+2^{j'}k+2^jk')} (-1 + e^{2^{-1-j'}\theta})^2 \\ & \times (1 - 2e^{2^{-1-j}\theta} + e^{2^{-j}\theta} + e^{2^{-j}(1+2k)\theta} + e^{2^{-j'}\theta(1+2k')} \\ & - 2e^{(1/2)\theta(2^{-j}+2^{1-j'}+2^{2-j'}k')}) \end{aligned}$$

3. *Negative containment case.*

$$\begin{aligned} \text{COV}(Y_{j,k}, Y_{j',k'}) &= \\ & \left(-\frac{\sigma^2}{\theta^3}\right) 2^{\frac{-2+j+j'}{2}} e^{-2^{-j-j'}\theta(2^j+2^{j'}+2^{j'}k+2^jk')} (-1 + e^{2^{-1-j'}\theta})^2 \\ & \times (1 - 2e^{2^{-1-j}\theta} + e^{2^{-j}\theta} + e^{2^{-j}(1+2k)\theta} - 2e^{2^{-1-j}(3+4k)\theta} + e^{2^{-j'}\theta(1+2k')}) \end{aligned}$$

4. *Identical support case.*

$$\begin{aligned} \text{COV}(Y_{j,k}, Y_{j',k'}) &= \mathbb{V}(Y_{j,k}) = \\ & \left(\frac{\sigma^2}{\theta^3}\right) 2^{-1+j} e^{-2^{1-j}(1+k)\theta} \\ & \times (-1 + 4e^{2^{-1-j}\theta} + 4e^{3(2^{-1-j})\theta} - e^{2^{1-j}\theta} - 6e^{2^{-j}\theta} - 2e^{2^{-j}(1+2k)\theta} \\ & + 8e^{2^{-1-j}(3+4k)\theta} + e^{2^{1-j}(1+k)\theta}(-6 + 2^{1-j}\theta)) \end{aligned}$$

Again, with these covariances, we can create an almost sure construction of the Ornstein–Uhlenbeck sample path, with the help of the Gram–Schmidt process.

5 Final remarks

We have shown that Brownian motion can be constructed almost surely with the converging sum:

$$B_t = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} Y_{j,k} h_{j,k}(t)$$

Similarly, for the Ornstein–Uhlenbeck process:

$$X_t = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} Y_{j,k} h_{j,k}(t)$$

In fact, the result can be extended to other Gaussian stochastic processes $X_t \in L^2([0, 1])$ (not to be confused with the Ornstein–Uhlenbeck process), where:

$$X_t = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} Y_{j,k} h_{j,k}(t) \quad (48)$$

and (49)

$$Y_{j,k} = \int_0^1 X_s h_{j,k}(s) ds$$

As long as the Gaussian stochastic process is in $L^2([0, 1])$, (49) holds from Theorem 1.2.5(1). The $Y_{j,k}$ are Gaussian random variables centered at 0, as shown in (46). Furthermore, our calculations of the covariances for $Y_{j,k}$ allow us to use the Gram–Schmidt process to construct the dependent $Y_{j,k}$ s using i.i.d. standard Gaussian random variables.

Further work is required to study how these results can be extended to other stochastic processes.

Appendices

A Zorn's Lemma

Definition A.0.1. A *partially ordered set* X is a set equipped with an order relation \leq , such that it satisfies the following three properties:

1. $x \leq x$ for every $x \in X$
2. If $x \leq y$ and $y \leq x$, then $x = y$, for $x, y \in X$
3. If $x \leq y$ and $y \leq z$, then $x \leq z$, for $x, y, z \in X$

Along with \leq , we will let $<$ define a strict relation such that if $x < y$ then $x \leq y$ and $x \neq y$.

Definition A.0.2. A subset Y of a partially ordered set X is *totally ordered* if for every $x, y \in Y$, either $x \leq y$ or $y \leq x$. Y is also known as a *chain*.

Lemma A.0.1. (Zorn's). *For every chain Y in a partially ordered set X , if there is some $u \in X$ such that $u \geq y$ for all $y \in Y$ (every Y has an upper bound), then X has a maximal element. An element $m \in X$ is a maximal element if $m \geq x$ for all $x \in X$.*

Proof. [6] Suppose X is partially ordered by \leq and every chain in X has an upper bound. By contradiction suppose X has no maximal element. Let us denote u as the upper bound of a chain C in X . Because X has no maximal element, we can choose an element $x \in X$ that is greater than u , such that $y < x$ for all $y \in C$. The element x is a *strict upper bound* of C . It follows from the axiom of choice that there exists a function f that maps all each chain $C \subseteq X$ to its strict upper bound $f(C) = x$.

Let us define the set P as $P(C, x) = \{y \in C \mid y < x\}$. We say that a subset of a chain C that has the form $P(C, x)$ is called an *initial segment* of C . Note that $P(C, x) \subseteq C$.

We define a subset A of X as *conforming* if the follow two conditions hold:

1. A is a chain every non-empty subset of A has a minimal element ($x \in A$ is a minimal element of A if $x \leq a$ for all $a \in A$).
2. For all x in A , $x = f(P(A, x))$

We claim that if A and B are conforming subsets of X and $A \neq B$, then one of these two sets is an initial segment of the other. To show the claim, assume that $A \setminus B \neq \emptyset$ (the case for $B \setminus A \neq \emptyset$ will yield the same result). We claim that $P(A, x) = B$.

To show that $P(A, x) \subseteq B$, we assume by contradiction that there is an element $a \in P(A, x)$ and $a \notin B$. By the definition of the set $P(A, x)$, $a < x$ and $a \in A$. However, if $a \in A$ and $a \notin B$, then $a \in A \setminus B$. But, $x \leq a$ which contradicts our choice of a .

To complete the proof we need to show that $B \subseteq P(A, x)$. By contradiction, suppose that $B \setminus P(A, x) \neq \emptyset$. Define y to be the minimal element of $B \setminus P(A, x)$. Following the same logic as above, $P(B, y) \subseteq P(A, x) \subseteq A$.

Note, consider any $u \in P(B, y)$ and any element $v \in A$ such that $v < u$. If $u \in P(B, y) \subseteq P(A, x)$, then $u < x$ because $u \in P(A, x)$ and thus $v < u < x$. Thus, $v \notin A \setminus B$ because x is the minimal element of $A \setminus B$. Thus, $v \in B$ and $v \in P(B, y)$ since $v < u < y$.

Define z to be the minimal element of $A \setminus P(B, y)$. It follows from the same logic that $P(A, z) \subseteq P(B, y)$. Note that our choice of z cannot be less than any element in $P(B, y)$, otherwise $z \in P(B, y)$ which contradicts $z \in A \setminus P(B, y)$. Thus, all elements of $P(B, y)$ are less than z (elements in $P(B, y)$ cannot equal z because $z \notin P(B, y)$). Furthermore, recall that $P(B, y) \subseteq A$. Thus, $P(B, y) \subseteq P(A, z)$. We conclude that $P(A, z) = P(B, y)$.

Since $x \in A \setminus B \subseteq A \setminus P(B, y)$, $x \in A \setminus P(B, y)$. Because z is the minimal element of $A \setminus P(B, y)$, it follows that $z \leq x$. By the definition of conforming sets, $z = f(P(A, z)) = f(P(B, y)) = y$. Since y is in B , z cannot equal x because x is defined to be in $A \setminus B$. Thus, $z < x$, and since $z \in A$, it follows that $y = z \in P(A, x)$. This contradicts the choice of y where $y \notin P(A, x)$. We conclude that $B \setminus P(A, x) = \emptyset$ and $P(A, x) = B$.

Note that conforming sets are chains by definition. From the above claim that if A and B are conforming subsets of X and $A \neq B$, then one of these two sets is an initial segment of the other, we observe that $B = P(A, x)$ for some $x \in A$. Therefore, $A \cup B = A$ and we conclude that the union of two conforming sets is conforming. Thus, the union, U , of all conforming subsets of X is conforming. Choose $x \in X$ such that $x = f(U)$ and note that $\{x\}$ is conforming. Thus, $U \cup \{x\}$ is conforming. Therefore, x is in U so x cannot be a strict upper bound of U . Thus, the chain U has no strict upper bound in X , contradicting the fact that X does not have a maximal element. □

B Fatou's Lemma

Lemma B.0.1. (Fatou's). *In the Lebesgue measure space, if $\{f_n\}$ is a sequence of integrable functions such that $f_n \geq 0$ a.e. for all n and $\liminf \int f_n d\lambda < \infty$, then $\liminf f_n$ is an integrable function and:*

$$\int \liminf f_n d\lambda \leq \liminf \int f_n d\lambda$$

Proof. [1] Define a function $g_n(x) = \inf\{f_i(x) \mid i \geq n\}$. Since $0 \leq g_n \leq f_n$, and g_n is measurable for all n , g_n is integrable. Since g_n is monotonically increasing and $\lim \int g_n d\lambda$ is bounded above by $\liminf \int f_n d\lambda < \infty$, it follows from Levi's Theorem that there exists an integrable function g such that $g_n \uparrow g$ a.e. Thus, $\lim g_n(x) = \liminf\{f_i(x) \mid i \geq n\}$, which gives us the infimum of the converging subsequence of $\{f_n(x)\}$. It follows that $g = \liminf f_n$ a.e. and thus $\liminf f_n$ is

integrable and:

$$\int \liminf f_n d\lambda = \int g d\lambda = \lim_{n \rightarrow \infty} \int g_n d\lambda \leq \liminf \int f_n d\lambda \quad (50)$$

□

C Fubini's Theorem

Lemma C.0.1. *Let (X, S, μ) and (Y, Σ, ν) be two fixed measure spaces and their product semiring $S \otimes \Sigma$ be:*

$$S \otimes \Sigma = \{A \times B \mid A \in S \text{ and } B \in \Sigma\}$$

Let $\mu \times \nu$ be a measure on $S \otimes \Sigma$ such that $\mu \times \nu = \mu(A)\nu(B)$. Furthermore, let $(\mu \times \nu)^ = \mu^*(A)\nu^*(B)$ be the outer measure on $S \otimes \Sigma$. Suppose E is a $\mu \times \nu$ -measurable subset of $X \times Y$ satisfying $(\mu \times \nu)^*(E) < \infty$. Then, for μ -almost all x the set E_x is a ν -measurable subset of Y , and there exists an integrable function $x \mapsto \nu^*(E_x)$ on X such that:*

$$(\mu \times \nu)^*(E) = \int_X \nu^*(E_x) d\mu(x) \quad (51)$$

Similarly, for ν -almost all y the set E^y is a μ -measurable subset of X , and there exists an integrable function $y \mapsto \mu^(E^y)$ on Y such that:*

$$(\mu \times \nu)^*(E) = \int_Y \mu^*(E^y) d\nu(y) \quad (52)$$

Proof. [1] This proof will have five steps.

Step 1. Let $E = A \times B \in S \otimes \Sigma$. Define a set E_x where $E_x = B$ if $x \in A$ and $E_x = \emptyset$ if $x \notin A$. Therefore, E_x is a ν -measurable subset of Y for all $x \in X$ and:

$$\nu(E_x) = \nu(B)\mathbb{1}_A(x) \quad (53)$$

Since $(\mu \times \nu)^*(E) = (\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B) < \infty$, there are two possible scenarios:

1. Both A and B have finite measure.

In this scenario, (53) shows that $x \mapsto \nu^*(E_x)$ is an integrable step function where:

$$\int_X \nu^*(E_x) d\mu(x) = \int \nu(B)\mathbb{1}_A d\mu = \mu(A) \cdot \nu(B) = (\mu \times \nu)^*(E) \quad (54)$$

2. Either A or B has infinite measure.

In this scenario, the other set must have measure 0. Thus, (53) guarantees that $\nu(E_x) = 0$ for μ -almost all x . Thus:

$$\int_X \nu^*(E_x) d\mu(x) = 0 = (\mu \times \nu)^*(E) \quad (55)$$

Step 2. Now, suppose that E is a σ -set (there exists a disjoint sequence of sets $\{E_n\}$ of $S \otimes \Sigma$ such that $E = \cup_{i=1}^{\infty} E_n$) of $S \otimes \Sigma$. By the definition of a σ -set, we can choose a disjoint sequence $\{E_n\}$ of $S \otimes \Sigma$ such that $E = \cup_{n=1}^{\infty} E_n$. Define $E_x = \cup_{n=1}^{\infty} (E_n)_x$. From step 1, it follows that E_x is a measurable subset of Y for all $x \in X$. Define a function $f(x) = \nu^*(E_x)$ and $f_n(x) = \sum_{i=1}^n \nu((E_i)_x)$. It follows from step 1 that each f_n represents an integrable function:

$$\int f_n d\mu = \sum_{i=1}^n \int_X \nu((E_i)_x) d\mu(x) = \sum_{i=1}^n (\mu \times \nu)(E_i) \uparrow (\mu \times \nu)^*(E) < \infty \quad (56)$$

Since the $(E_n)_x$ are disjoint in Σ , $\nu^*(E_x) = \sum_{n=1}^{\infty} \nu((E_n)_x)$. Thus, $f_n(x) \uparrow f(x)$. It follows from Levi's Theorem that f is an integrable function and:

$$\int_X \nu^*(E_x) d\mu(x) = \int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu = \sum_{i=1}^{\infty} (\mu \times \nu)(E_i) = (\mu \times \nu)^*(E) \quad (57)$$

Step 3. Now, suppose that E is a countable intersection of σ -sets of finite measure. Choose a sequence $\{E_n\}$ of σ -sets such that $E = \cap_{n=1}^{\infty} E_n$, $(\mu \times \nu)^*(E_1) < \infty$ and $E_{n+1} \subseteq E_n$ holds for all n . For all n , define a function $g_n(x)$ such that $g_n(x) = 0$ if $\nu^*((E_n)_x) = \infty$ and $g_n(x) = \nu^*((E_n)_x)$ if $\nu^*((E_n)_x) < \infty$. From step 2, each g_n is an integrable function over X such that $\int g_n d\mu = (\mu \times \nu)^*(E_n)$. Define $E_x = \cap_{n=1}^{\infty} (E_n)_x$, again E_x is a ν -measurable set for all x in X . Since $\nu^*((E_n)_x) < \infty$ for μ -almost all x , it follows that $g_n(x) = \nu^*((E_n)_x) \downarrow \nu^*(E_x)$ for μ -almost all x . Thus, $x \mapsto \nu^*(E_x)$ is an integrable function and:

$$\int_X \nu^*(E_x) d\mu(x) = \lim_{n \rightarrow \infty} \int g_n d\mu = \lim_{n \rightarrow \infty} (\mu \times \nu)^*(E_n) = (\mu \times \nu)^*(E) \quad (58)$$

Step 4. Now, suppose that E is a null set. Therefore, $(\mu \times \nu)^*(E) = 0$. For all i , choose a sequence $\{E_n^i\}$ such that $E \subseteq \cup_{n=1}^{\infty} E_n^i$ and $\sum_{n=1}^{\infty} (\mu \times \nu)(E_n^i) < (\mu \times \nu)^*(E) + \frac{1}{i}$. Define a set $G_i = \cup_{n=1}^{\infty} E_n^i$. Then all G_i are measurable sets and $E \subseteq G_i$. Let $G = \cap_{n=1}^{\infty} G_n$. Thus, $E \subseteq G$ and G is measurable. For all i :

$$\mu^*(E) \leq (\mu \times \nu)^*(G) \leq (\mu \times \nu)^*(G_i) \leq \sum_{n=1}^{\infty} (\mu \times \nu)(E_n^i) < (\mu \times \nu)^*(E) + \frac{1}{i} \quad (59)$$

It follows that $(\mu \times \nu)^*(G) = (\mu \times \nu)^*(E)$. From step 3:

$$\int_X \nu^*(G_x) d\mu(x) = (\mu \times \nu)^*(G) = 0. \quad (60)$$

Thus, $\nu^*(G_x) = 0$ holds for μ -almost all x . Since $E_x \subseteq G_x$, it follows that $\nu^*(E_x) = 0$ for μ -almost all x . Thus, E_x is ν -measurable for μ -almost all x and:

$$\int_X \nu^*(E_x) d\mu(x) = 0 = (\mu \times \nu)^*(E) \quad (61)$$

Step 5. Choose a $\mu \times \nu$ -measurable set F that is a countable intersection of σ -sets with finite measure such that $E \subseteq F$ and $(\mu \times \nu)^*(F) = (\mu \times \nu)^*(E)$. Let $G = F \setminus E$. Therefore, G is the null set and from step 4, $\nu^*(G_x) = 0$ for μ -almost all x . From step 3, $x \mapsto \nu^*(F_x)$ is an integrable function and thus $x \mapsto \nu^*(E_x)$ is also an integrable function such that:

$$(\mu \times \nu)^*(E) = (\mu \times \nu)^*(F) = \int_X \nu^*(F_x) d\mu(x) = \int_X \nu^*(E_x) d\mu(x) \quad (62)$$

Thus, it has been shown that $(\mu \times \nu)^*(E) = \int_X \nu^*(E_x) d\mu(x)$. (52) can be shown in a similar fashion as above. \square

Theorem C.0.1. (Fubini's). *Let (X, S, μ) and (Y, Σ, ν) be two fixed measure spaces. If $f : (X \times Y) \mapsto \mathbb{R}$ is a $\mu \times \nu$ -integrable function, then:*

$$\int_{X \times Y} f d(\mu \times \nu) = \int_Y \int_X f d\mu d\nu = \int_X \int_Y f d\nu d\mu$$

Proof. [1] It is sufficient to prove this for the case that $f(x, y) \geq 0$ since we can express f as $f^+ - f^-$. Consider a sequence of simple functions $\{s_n\}$ such that $0 \leq s_n(x, y) \uparrow f(x, y)$ holds for all x, y . It follows that:

$$\int_X \left[\int_Y s_n(x, y) d\nu(y) \right] d\mu(x) = \int_{X \times Y} s_n d(\mu \times \nu) \uparrow \int_{X \times Y} f d(\mu \times \nu) < \infty \quad (63)$$

From Lemma C.0.1, we can define an integrable function $g_n(x)$ over X :

$$g_n(x) = \int (s_n)_x d\nu = \int_Y s_n(x, y) d\nu(y) \quad (64)$$

Note that $g_n(x)$ monotonically increases for μ -almost all x . From Levi's Theorem, there exists a μ -integrable function $g : X \mapsto \mathbb{R}$ such that $g_n(x) \uparrow g(x)$ holds for μ -a.e. It follows from (64) that $\int (s_n)_x d\nu \uparrow g(x) < \infty$. Since $(s_n)_x \uparrow f_x$ holds for all x , it follows that f_x is ν -integrable a.e. over X and:

$$g_n(x) = \int (s_n)_x d\nu = \int_Y s_n(x, y) d\nu(y) \uparrow \int_Y f_x d\nu \quad (65)$$

It then follows from (63) and (65) that the function g is an integrable function such that:

$$\int f d(\mu \times \nu) = \int_X \left(\int_Y f_x d\nu \right) d\mu = \int \int f d\nu d\mu \quad (66)$$

In a similar fashion as above, it can be shown that:

$$\int f d(\mu \times \nu) = \int_Y \left(\int_X f_x d\mu \right) d\nu = \int \int f d\mu d\nu \quad (67)$$

\square

D Brownian motion simulation

An application of our result is that we can create a simulation of a Brownian motion sample path. Specifically, we can generate a sequence of i.i.d. standard Gaussian random variables and construct a sequence of $\{Y_{j,k}\}$ Gaussian random variables via the Gram–Schmidt process, then sum the product $Y_{j,k}h_{j,k}(t)$ up to our desired dyadic level.

Below is a simulation done in R, up to the third dyadic level.

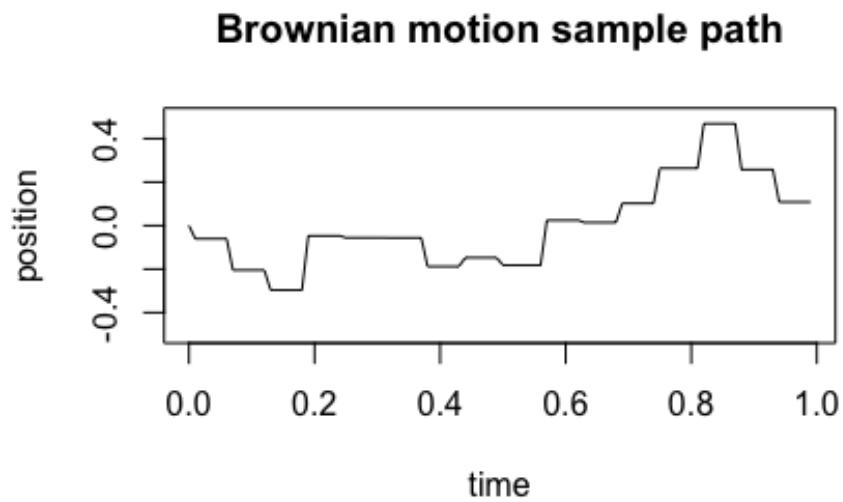


Figure 2: [8] Haar expansion of Brownian motion up to the third dyadic level.

References

- [1] C.D. Aliprantis, O. Burkinshaw, *Principles of Real Analysis*, Academic Press, 1998.
- [2] D. Walnut. 2017: Lecture 6 - Orthonormal Wavelet Bases, http://www1.chapman.edu/~mbvajiac/conferences/chapman_lec06_slides.pdf (2017)
- [3] H.L. Royden, P.M. Fitzpatrick, *Real Analysis*, Pearson Education, 2010.
- [4] J. Belk. Appendix: Sum Over Sets <http://faculty.bard.edu/belk/math461/SumsOverSets.pdf> (2016)
- [5] J. Bell. Khinchin's inequality and Etemadi's inequality <http://individual.utoronto.ca/jordanbell/notes/khinchin.pdf> (2015)
- [6] J. Lewin, J.: A Simple Proof of Zorn's Lemma, *The American Mathematical Monthly*, **98**, 353-354 (1991).
- [7] P. Billingsley, *Probability and Measure*, John Wiley & Sons, Inc., New York, 1995.
- [8] S. Tindel. Brownian motion <https://www.math.purdue.edu/~stindel/teaching/ma539/brownian-motion2.pdf> (2018)