ITÔ'S FORMULA WITH APPLICATIONS

ETHAN NAEGELE

Abstract. In this expository paper, we highlight some of the far-reaching applications of Itô's Formula, a powerful tool from stochastic calculus which can give quick, enlightening proofs of robust results in probability and other areas of analysis. We build stochastic integration, then prove Itô's Formula before discussing some of its applications in probability, partial differential equations, and complex analysis.

Contents

1. Introduction

Brownian motion is a central object in probability. A particularly rich theory involving Brownian motion is that of stochastic integration, in which one integrates random processes with respect to Brownian motion. Defining such an integral is itself a challenge, because, as we will see, one cannot define a stochastic integral in the usual Lebesgue-Stieltjes sense. Much of our effort, therefore, will be in defining the stochastic integral.

Beyond the mere definition, however, lies the question of how the stochastic integral can be useful. This paper gives a glimpse into some of the applications of the stochastic integral primarily through applying a result known as Itô's formula, which is effectively the fundamental theorem of stochastic calculus. Itô's formula is a powerful result, having applications in areas such as partial differential equations and complex analysis. Such applications include proving a Feynman-Kac formula (Theorem 4.3) and proving conformal invariance of Brownian motion (Theorem 4.10), of which Liouville's theorem from complex analysis is a corollary (Theorem 4.12).

To prove these results with Itô's formula, we begin by collecting the necessary results about Brownian motion in Section 2. We then construct the Itô integral

Date: August 29, 2023.

and prove Itô's formula in Section 3, and we see what this powerful result can do in Section 4.

We assume the reader is familiar with the basics of measure-theoretic probability, but we will remind the reader of the necessary details throughout. Of special importance is understanding the different types of convergence and their relations.

2. Brownian Motion

We begin by introducing Brownian motion, the underlying object that will be central to our endeavors, first considering the one-dimensional case.

Definition 2.1. A linear (i.e., one-dimensional) ***Brownian motion*** started at $x \in \mathbb{R}$ is a real-valued stochastic process $(B_t)_{t\geq 0}$ satisfing the following properties:

- $B_0 = x$.
- Independent increments: For all times $0 \leq s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq s_n \leq t_n$, the increments $B_{t_1} - B_{s_1}, ..., B_{t_n} - B_{s_n}$ are independent random variables.
- Increment distribution: For any $t \geq 0$ and $h > 0$, the increment $B(t+h) - B(t)$ is a $N(0,h)$ random variable.
- The map $t \mapsto B_t$ is almost surely continuous.

If $x = 0$, we call $(B_t)_{t\geq 0}$ a ***standard Brownian motion***. For dimension $d \geq 2$, we define ***d-dimensional Brownian motion*** to be a process $(B_t)_{t\geq 0}$ where $B(t) = (B_t^1, ..., B_t^d)$ such that each component is an independent linear Brownian motion, where independence of processes is in the sense defined below.

Definition 2.2. We call two processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ ***independent*** if for any times $t_1, ..., t_n \geq 0$ and $s_1, ..., s_m \geq 0$, the vectors $(X_{t_1}, ..., X_{t_n})$ and $(Y_{t_1}, ..., Y_{t_m})$ are independent. Similarly, a process $(X_t)_{t\geq 0}$ is ***independent*** of a σ-algebra $\mathcal{A}$ if $\sigma(X_t : t \geq 0)$ is independent of $\mathcal{A}$.

We have defined Brownian motion, but it must be shown to exist. We will not describe this construction for sake of brevity. The proof is provided in [2, Theorem 1.3].

Theorem 2.3. *Standard Brownian motion exists.*

We will now collect some facts about the regularity of Brownian motion that will help us throughout our endeavors. To do so, we first recall a definition from analysis.

Definition 2.4. A function $f : [0, \infty) \to \mathbb{R}$ is ***locally α-Hölder continuous*** at x if there exists $\delta > 0$ and $C > 0$ such that

$$|f(x) - f(y)| \leq C|x - y|^{\alpha}$$

for all $y \geq 0$ with $|x - y| < \delta$.

We now have the following statement on the Hölder regularity of Brownian motion as appearing in [2, Corollary 1.20].

Proposition 2.5. *If $\alpha < 1/2$, then, almost surely, Brownian motion is locally α-Hölder continuous at every $x \in \mathbb{R}$.*

To lay the groundwork for our discussion of stochastic calculus in Section 3, we must discuss a fundamental concept known as quadratic variation.

Recall from analysis that a function $f : [0, t] \to \mathbb{R}$ is of bounded variation if

$$\sup \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| < \infty,$$

where the sup is taken over all $n \in \mathbb{N}$ and partitions $0 = t_0 < t_1 < \cdots < t_n = t$. Functions of bounded variation are in some sense "nice" in that they are, for example, differentiable almost everywhere. More importantly for our purposes, one can construct a Lebesgue-Stieltjes integral with respect to f if f is of bounded variation. As one might expect, we are not so fortunate to have encountered such a "nice" function with Brownian motion, as we shall see. However, the notion of quadratic variation will save us in our endeavors to create a stochastic integral later on.

Definition 2.6. Consider a sequence $(\mathcal{P}_n)_{n\geq 1}$ of partitions of $[0, t]$, where $\mathcal{P}_n$ is of the form

$$0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = t.$$

Define the ***mesh size*** of the nth partition to be

$$\Delta(n) := \sup_{1\leq i\leq k_n} (t_i^n - t_{i-1}^n).$$

Define the ***quadratic variation*** of linear Brownian motion to be

$$[B]_t := \lim_{\Delta(n)\to 0} \sum_{i=1}^{k_n} (B_{t_i^n} - B_{t_{i-1}^n})^2,$$

where the limit is in probability.

More generally, define the ***covariation*** of two continuous processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ to be

$$[X, Y]_t := \lim_{\Delta(n)\to 0} \sum_{i=1}^{k_n} (X_{t_i^n} - X_{t_{i-1}^n})(Y_{t_i^n} - Y_{t_{i-1}^n}),$$

where the limit is in probability, provided it exists.

Theorem 2.7 shows that quadratic variation of Brownian motion exists in probability, so our definition is valid. Observe that covariation behaves algebraically similar to covariance, in that by our definitions, $[X, X]_t = [X]_t$ and covariation is bilinear.

Let's try to apply these definitions to Brownian motion. First, with all of the erratic fluctuations that a Brownian path makes, no matter how far one "zooms in" due to the self-similarity of the path, it seems plausible that these fluctuations constitute a path of unbounded variation. We can intuitively reason that after some time ε has passed, we would expect roughly that Brownian motion has moved a distance of $\sqrt{\varepsilon}$, because $\mathbf{E}[B_\varepsilon^2] = \varepsilon$. For simplicity, considering Brownian motion on $I = [0, 1]$, there are roughly $1/\varepsilon$ subintervals of length ε on I, and, since on each subinterval $[t_i, t_{i+1}]$, we have $|B_{t_{i+1}} - B_{t_i}| \approx \sqrt{\varepsilon}$, we have that $\sum |B_{t_{i+1}} - B_{t_i}|$ is roughly $\sqrt{\varepsilon}(1/\varepsilon) = 1/\sqrt{\varepsilon}$. It is now clear that the latter diverges as $\varepsilon \to 0$.

When we instead consider quadratic variation, on these same subintervals, we now have $(B_{t_{i+1}} - B_{t_i})^2 \approx \varepsilon$, suggesting that $\sum (B_{t_{i+1}} - B_{t_i})^2$ is roughly $\varepsilon(1/\varepsilon) = 1$, which suggests that we would have $[B]_1 = 1$, and in general $[B]_t = t$. Our intuition here is correct and is formalized in the theorem below, whose proof comes from [5, Theorem 6.12.1] and [2, Theorem 1.35].

Theorem 2.7. *(1) In probability, we have $[B]_t = t$ for all $t > 0$.*

(2) If, in addition, we assume that $\lim_{n\to\infty} n^2\Delta(n) = 0$, then we have $[B]_t = t$ almost surely. Therefore, Brownian motion is almost surely of unbounded variation.

(3) Alternatively, if $(\mathcal{P}_n)_{n\geq 1}$ is a nested sequence of partitions, that is, if for any n, we have $\mathcal{P}_n \subset \mathcal{P}_{n+1}$, then, again, $[B]_t = t$ almost surely.

Proof. (1) It is enough to prove L^2 convergence. That is, we will show

$$\lim_{n\to\infty} \mathbf{E}\left[\left(\sum_{i=1}^{k_n}(B_{t_i^n} - B_{t_{i-1}^n})^2 - t\right)^2\right] = 0.$$

A fact we will use is that for a random variable $Y \sim N(0, \sigma^2)$, we have $\mathbf{E}[Y^4] = 3\sigma^4$. Observe that the random variables $X_i := (B_{t_i^n} - B_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)$ are i.i.d. of mean 0, hence $\mathrm{Var}[X_i] = \mathbf{E}[X_i^2]$, and independence implies $\mathrm{Var}[\sum_{i=1}^{k_n} X_i] = \sum_{i=1}^{k_n}\mathrm{Var}[X_i]$. It then follows that

$$\begin{aligned}
\mathbf{E}\left[\left(\sum_{i=1}^{k_n}[(B_{t_i^n} - B_{t_{i-1}^n})^2] - t\right)^2\right] &= \mathbf{E}\left[\left(\sum_{i=1}^{k_n} X_i\right)^2\right] = \sum_{i=1}^{k_n}\mathbf{E}[X_i^2] \\
&= \sum_{i=1}^{k_n}\mathbf{E}[(B_{t_i^n} - B_{t_{i-1}^n})^4] - 2\sum_{i=1}^{k_n}\mathbf{E}[(B_{t_i^n} - B_{t_{i-1}^n})^2](t_i^n - t_{i-1}^n) + \sum_{i=1}^{k_n}(t_i^n - t_{i-1}^n)^2 \\
&= 3\sum_{i=1}^{k_n}(t_i^n - t_{i-1}^n)^2 - 2\sum_{i=1}^{k_n}(t_i^n - t_{i-1}^n)^2 + \sum_{i=1}^{k_n}(t_i^n - t_{i-1}^n)^2 \\
&\leq 2\Delta(n)\sum_{i=1}^{k_n}(t_i^n - t_{i-1}^n) \\
&= 2\Delta(n)t,
\end{aligned}$$

which tends to 0 and proves the first statement.

(2) The assumption implies that $\Delta(n) = x_n/n^2$ where $x_n \to 0$. Markov's inequality implies that

$$\begin{aligned}
\mathbf{P}\left[\left(\sum_{i=1}^{k_n}(B_{t_i^n} - B_{t_{i-1}^n})^2 - t\right)^2 > 2x_n\right] &\leq \mathbf{E}\left[\left(\sum_{i=1}^{k_n}(B_{t_i^n} - B_{t_{i-1}^n})^2 - t\right)^2\right]/2x_n \\
&\leq 2\Delta(n)t/2x_n \\
&= t/n^2.
\end{aligned}$$

Therefore, the Borel-Cantelli lemma implies that there exists some N such that $n \geq N$ implies $\mathbf{P}\left[\left(\sum_{i=1}^{k_n}(B_{t_i^n} - B_{t_{i-1}^n})^2 - t\right)^2 \leq 2x_n\right] = 1$, hence taking $n \to \infty$ inside the probability proves almost sure convergence.

To see that Brownian motion is of unbounded variation, let $\alpha \in (0, 1/2), C > 0$, and $\delta > 0$ as in the statement of Proposition 2.5 applied to the interval $[0, t]$. Now, take n so large that for each $t_i^n, t_{i-1}^n \in \mathcal{P}_n$, we have $|t_i^n - t_{i-1}^n| \le \Delta(n) < \delta$. Then it follows that

$$\sum_{i=1}^{k_n} |B_{t_i^n} - B_{t_{i-1}^n}| \ge (C\Delta(n)^\alpha)^{-1} \sum_{i=1}^{k_n} (B_{t_i^n} - B_{t_{i-1}^n})^2.$$

Taking $\Delta(n) \to 0$, the limit on the right is almost surely infinite, hence the left is too. It follows that Brownian motion is of unbounded variation almost surely.

(3) The proof of this statement is found in [2, Theorem 1.35]. It uses the theory of reverse martingales, which we will not discuss here. □

Definition 2.8. We define the Brownian motion's ***natural filtration*** as follows:

$$\mathcal{F}^0(t) := \sigma(B_s : 0 \le s \le t).$$

Additionally, we define the ***augmented filtration***:

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

Definition 2.9. If $(\Omega, \mathcal{A}, (\mathcal{F}(t))_{t\ge 0}, \mathbf{P})$ is a filtered probability space, T is a ***stopping time*** with respect to $(\mathcal{F}(t))_{t\ge 0}$ if $\{T \le t\} \in \mathcal{F}(t)$ for all $t \ge 0$, in which case we can define

$$\mathcal{F}(T) := \{A \in \mathcal{A} : A \cap \{T \le t\} \in \mathcal{F}(t) \text{ for all } t \ge 0\}.$$

Intuitively, the natural filtration contains the information of Brownian motion up to time t (all events $\{B_s \in B\}$, where $0 \le s \le t$ and B is a Borel set). The augmented filtration is somewhat of an upgrade compared to the natural; it is clear that $\mathcal{F}^0(t) \subset \mathcal{F}^+(t)$, and it contains events that depend on times that are arbitrarily soon after t. In this way, one can think of $\mathcal{F}^+(t)$ as giving an "infinitesimal peek" past t.

In general, we will use the augmented filtration instead of the natural. The augmented strictly contains all of the stopping times of the natural. For instance, hitting times of open sets are stopping times with respect to $\mathcal{F}^+$, but not to $\mathcal{F}^0$ in general (details omitted).

Theorem 2.10 (Markov Property)**.** *Let $(B_t)_{t\ge 0}$ be a d-dimensional Brownian motion started at $x \in \mathbb{R}^d$. Then, if $s > 0$, the process $(B_{t+s} - B_s)_{t\ge 0}$ is a Brownian motion started at the origin which is independent of $(B_t)_{0\le t\le s}$.*

Proof. One observes that all the properties stated in Definition 2.1 hold. The independence of these processes follows from the independence of the increments of Brownian motion. To see this, first consider the start point $x = 0$. Then by independent increments, $(B_{t_1+s} - B_s, ..., B_{t_n+s} - B_s)$ and $(B_{s_1}, ..., B_{s_m})$ are independent vectors because $B_{s_i} = B_{s_i} - B_0$. Then, for an arbitrary start point $x \in \mathbb{R}^d$, adding the constant vector x to the second vector does not change the independence of the two vectors. □

The Markov property essentially tells us that after time s, the Brownian motion is "refreshed" and behaves like a new Brownian motion that starts at B_s, independent of whatever happened before time s. Another version of Theorem 2.10 is true for the augmented filtration.

Theorem 2.11. *If $(B_t)_{t\geq 0}$ is a Brownian motion started in $x \in \mathbb{R}^d$ and $s \geq 0$, the process $(B_{t+s} - B_s)_{t\geq 0}$ is a Brownian motion started at 0 and independent of $\mathcal{F}^+(s)$.*

To appreciate the next example, we recall that a ***continuous-time martingale*** with respect to a filtration $(\mathcal{F}(t))_{t\geq 0}$ is a process $(X_t)_{t\geq 0}$ such that (X_t) is adapted to the filtration (that is, X_t is $\mathcal{F}(t)$-measurable for every $t \geq 0$), $\mathbf{E}|X_t| < \infty$ for all $t \geq 0$, and for all $s < t$, we have $\mathbf{E}[X_t \mid \mathcal{F}(s)] = X_s$.

Example 2.12. Brownian motion is a martingale with respect to $(\mathcal{F}^+(t))_{t\geq 0}$ since, by Theorem 2.11, for any $s < t$ we have

$$\begin{aligned}\mathbf{E}[B_t \mid \mathcal{F}^+(s)] &= \mathbf{E}[B_t - B_s + B_s \mid \mathcal{F}^+(s)] \\ &= \mathbf{E}[B_t - B_s \mid \mathcal{F}^+(s)] + B_s \\ &= \mathbf{E}[B_t - B_s] + B_s \\ &= B_s.\end{aligned}$$

Example 2.13. Let $s < t$. Then

$$\begin{aligned}\mathrm{Cov}(B_s, B_t) &= \mathbf{E}[B_s B_t] - \mathbf{E}[B_s]\mathbf{E}[B_t] \\ &= \mathbf{E}[B_s(B_t - B_s) + B_s^2] \\ &= \mathbf{E}[B_s]\mathbf{E}[B_t - B_s] + \mathbf{E}[B_s^2] \\ &= s.\end{aligned}$$

Now, Theorem 2.10 receives another upgrade with the Strong Markov property, stating that the Markov property extends to stopping times. The next two results are found in Theorem 2.16 and Theorem 2.49 respectively in [2].

Theorem 2.14 (Strong Markov Property)**.** *Let T be an almost surely finite stopping time with respect to $\mathcal{F}^+$. Then the process $(B_{t+T} - B_T)_{t\geq 0}$ is a standard Brownian motion independent of $\mathcal{F}^+(T)$.*

The following is a consequence of the Optional Stopping Theorem for continuous martingales (see [2, Proposition 2.42]).

Proposition 2.15 (Gambler's Ruin)**.** *Let (B_t) be a standard linear Brownian motion and $a < 0 < b$. Let $T := \min\{t \geq 0 : B_t \in \{a, b\}\}$. Then*

$$\mathbf{P}[B_T = a] = \frac{b}{|a| + b}, \ \mathbf{P}[B_T = b] = \frac{|a|}{|a| + b}, \ \text{and } \mathbf{E}[T] = |a|b.$$

To appreciate the following result, which is stated and proven in [7, Theorem 2.2.5], we note that Brownian motion is called recurrent if, loosely speaking, it "comes back infinitely often" and transient if it does not, having an infinite limit almost surely.

The result below states that Brownian motion in the plane is neighborhood recurrent, meaning it returns to neighborhoods infinitely often. This is a special result in that recurrence of Brownian motion does not hold in higher dimensions (see [2, Theorem 3.20]).

Theorem 2.16. *Let $U \subset \mathbb{R}^2$ be open. Then the set $\{t \geq 0 : B_t \in U\}$ is unbounded almost surely.*

3. Stochastic Integration and Itô's Formula

3.1. **The Itô Integral.** What would it mean to integrate with respect to Brownian motion? Let's take a look at a discrete analogue of the situation, where we have a simple random walk (S_n), where each increment $S_{n+1} - S_n$ is a step one unit to the left or right with equal probability. If we were to integrate a function H_n with respect to S_n, we would have something like

$$\int_0^n H_i dS_i = \sum_{i=0}^{n-1} H_{i+1}(S_{i+1} - S_i).$$

If we were to imagine a gambler betting on each step of the random walk, $\int_0^n H_i dS_i$ denotes the amount of winnings the gambler has at time n betting H_{i+1} on the walk moving to the right at time $i+1$, or betting $-H_{i+1}$ on the walk moving to the left. In this interpretation, we see that we would not want to have $H_{i+1} \in \mathcal{F}_{i+1}$, meaning the gambler could change his bet based on present information, so we should instead have $H_{i+1} \in \mathcal{F}_i$. This intuition is confirmed when we define the stochastic integral of step processes later. Overall, then, we gain the sense that our stochastic integral would roughly be like placing bets, specified by our integrand H, on infinitesimal increments $B_{t_{i+1}} - B_{t_i}$ of Brownian motion.

We would then like to integrate with respect to Brownian motion. We would imagine having some limit of the form $\lim \sum_i H_{x_i}(B_{t_{i+1}} - B_{t_i})$ where $t_1, ..., t_{n+1}$ is a partition of $[0,t]$, as in the usual case of Lebesgue-Steiltjes integration. Indeed, we have the following result from analysis (see [4, p. 41], [3, pp. 316-324]) highlighting our ideal formula that we would like to imitate with our stochastic integral:

Theorem 3.1. *Let A be a random process of almost surely bounded variation on every compact interval of $\mathbb{R}_{\geq 0}$, and let H be a jointly measurable function such that $s \mapsto H(s,\omega)$ is continuous for almost all ω. Let $(\mathcal{P}_n)$ be a sequence of finite random partitions of $[0,t]$ with $\lim_{n\to\infty} \Delta(n) = 0$. Then for $x_i \in [t_i, t_{i+1}]$ (with the $t_j \in \mathcal{P}_n$), we have almost surely*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} H_{x_i}(A_{t_{i+1}} - A_{t_i}) = \int_0^t H_s dA_s.$$

But there is a problem. As we saw in Theorem 2.7, Brownian motion is almost surely of *unbounded variation*. Is all hope lost? Indeed, according to the following proposition from [4, pp. 43], whose proof we omit, it is not possible to define an integral with respect to Brownian motion as a usual Stieltjes integral.

Proposition 3.2. *If $(\mathcal{P}_n)$ is a nested sequence of partitions, $x(t)$ is right continuous on $[0,1]$, and the sums $\sum_{i=1}^{k_n} h(t_i)(x(t_{i+1}) - x(t_i))$ converge to a limit for every continuous function h, then x is of bounded variation.*

It is clear, then, that we need to try a different approach. In particular, let's try to go for a weaker form of limit, the L^2 limit. Recall that (X_n) converges to X in the L^2 sense if $\lim_{n\to\infty} \mathbf{E}[(X_n - X)^2] = 0$.

In the L^2 theory of integration that we will therefore consider, the class of potential integrands are the so-called progressively measurable functions, which we define below. Before doing so, we note that from now on, we assume the probability space has a filtration $(\mathcal{F}(t))_{t\geq 0}$ to which Brownian motion is adapted, contains all events of probability 0, and such that the Strong Markov Property holds.

Definition 3.3. A process $(X_t(\omega))_{t\geq 0,\omega\in\Omega}$ is ***progressively measurable*** if for any $t \geq 0$, the map $X : [0,t] \times \Omega \to \mathbb{R}$ is measurable with respect to $\mathcal{B}([0,t]) \otimes \mathcal{F}(t)$.

Progressively measurable processes, roughly speaking, require only information from the Brownian path up to time t to determine the values of the process up to time t; they do not require information about Brownian motion from the future. The lemma from [2, Lemma 7.2], whose proof we omit, gives a large class of progressively measurable processes.

Lemma 3.4. *Processes $(X_t)_{t\geq 0}$ that are left or right continuous are also progressively measurable.*

We will now define the stochastic integral for step processes, which has a natural definition, and then we will build the general stochastic integral from there.

Definition 3.5. Consider a partition $0 = t_0 < t_1 < \cdots < t_{n+1} = t$ of $[0,t]$. A progressively measurable ***step process*** $(H_t(\omega))_{t\geq 0,\omega\in\Omega}$ is a function of the form

$$H_t(\omega) := \sum_{i=1}^{n} A_i(\omega)\mathbf{1}_{(t_i,t_{i+1}]}(t)$$

where A_i is a $\mathcal{F}(t_i)$-measurable random variable.

We define the ***stochastic integral of a step process*** H to therefore be

$$\int_0^\infty H_s dB_s = \sum_{i=1}^{n} A_i(B_{t_{i+1}} - B_{t_i}).$$

Having defined the stochastic integral of step processes, we now define the stochastic integral in general for progressively measurable functions.

Definition 3.6. Define the norm $\|H\|_2 := (\mathbf{E}\left[\int_0^\infty H_s^2 ds\right])^{1/2}$. If H is a progressively measurable process with $\|H\|_2 < \infty$ and (H_n) is a sequence of progressively measurable step processes such that $\|H_n - H\|_2 \to 0$, we define the ***stochastic integral*** of H to be

$$\int_0^\infty H_s dB_s := \lim_{n\to\infty} \int_0^\infty H_n(s) dB_s,$$

where the limit is in the L^2 sense.

In Definition 3.5, A_i's being $\mathcal{F}(t_i)$-measurable means that Brownian motion's information up to the beginning of the interval $(t_i, t_{i+1}]$ is enough to determine the value of H on that entire interval.

To make sense of Definition 3.6, we need to know that any progressively measurable function H can be approximated by such a sequence (H_n), that the L^2 limit indeed exists, and that it does not depend on the choice of approximating sequence (H_n). These steps are done in Lemma 3.7–Theorem 3.10, the first of which we state below.

Lemma 3.7. *Let $(H_s(\omega))_{s\geq 0,\omega\in\Omega}$ be a progressively measurable process such that $\|H\|_2 < \infty$. Then there exists a sequence of progressively measurable step processes (H_n) such that* $\lim_{n\to\infty}\|H_n - H\|_2 = 0$.

The approximation process is rather tedious and long, and is omitted but is in [2, Lemma 7.3].

With the approximation lemma now in hand, we turn to our next crucial tool, which establishes that the usual L^2 norm of the stochastic integral for step processes coincides with the $\|\cdot\|_2$ norm that we have established. This will lead to proving the isometry in general.

Lemma 3.8 (Itô Isometry). *Let H be a progressively measurable step process with $\mathbf{E}[\int_0^\infty H_s^2 ds] < \infty$. Then*

$$\mathbf{E}\left[\left(\int_0^\infty H_s dB_s\right)^2\right] = \mathbf{E}\left[\int_0^\infty H_s^2 ds\right].$$

Proof. If $i < j$, then

$$\begin{aligned}\mathbf{E}[A_iA_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})] &= \mathbf{E}[\mathbf{E}[A_iA_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) \mid \mathcal{F}(t_j)]] \\ &= \mathbf{E}[A_iA_j(B_{t_{i+1}} - B_{t_i})\mathbf{E}[(B_{t_{j+1}} - B_{t_j})]] \\ &= 0,\end{aligned}$$

where in moving from the first line to the second we used that A_i and A_j are both $\mathcal{F}(t_j)$-measurable, and the conditioning is removed from the expectation by Theorem 2.11. Thus,

$$\begin{aligned}\mathbf{E}\left[\left(\int_0^\infty H_s dB_s\right)^2\right] &= \mathbf{E}\left[\left(\sum_{i=1}^n A_i(B_{t_{i+1}} - B_{t_i})\right)^2\right] \\ &= \sum_{i=1}^n \mathbf{E}[A_i^2(B_{t_{i+1}} - B_{t_i})^2] \\ &= \sum_{i=1}^n \mathbf{E}[\mathbf{E}[A_i^2(B_{t_{i+1}} - B_{t_i})^2 \mid \mathcal{F}(t_i)]] \\ &= \sum_{i=1}^n \mathbf{E}[A_i^2\mathbf{E}[(B_{t_{i+1}} - B_{t_i})^2]] \\ &= \sum_{i=1}^n \mathbf{E}[A_i^2](t_{i+1} - t_i),\end{aligned}$$

where the fourth line follows from the third by using similar logic as above. Now,

$$\begin{aligned}\mathbf{E}\left[\int_0^\infty H_s^2 ds\right] &= \mathbf{E}\left[\int_0^\infty \left(\sum_{i=1}^n A_i\mathbf{1}_{(t_i,t_{i+1}]}\right)^2 ds\right] \\ &= \mathbf{E}\left[\int_0^\infty \left(\sum_{i=1}^n A_i^2\mathbf{1}_{(t_i,t_{i+1}]}^2 + 2\sum_{i<j} A_iA_j\mathbf{1}_{(t_i,t_{i+1}]}\mathbf{1}_{(t_j,t_{j+1}]}\right) ds\right] \\ &= \mathbf{E}\left[\int_0^\infty \left(\sum_{i=1}^n A_i^2\mathbf{1}_{(t_i,t_{i+1}]}\right) ds\right] \\ &= \mathbf{E}\left[\sum_{i=1}^n A_i^2(t_{i+1} - t_i)\right],\end{aligned}$$

where the elimination of the cross term follows from the disjointness of the intervals $(t_i, t_{i+1}]$ and $(t_j, t_{j+1}]$. Applying linearity of expectation to the last line gives the statement. □

Lemma 3.9. *Suppose* (H_n) *is a sequence of progressively measurable step processes such that*

$$\lim_{n,m\to\infty} \mathbf{E}\left[\int_0^\infty (H_n(s) - H_m(s))^2 ds\right] = 0.$$

Then

$$\lim_{n,m\to\infty} \mathbf{E}\left[\left(\int_0^\infty H_n(s) - H_m(s) dB_s\right)^2\right] = 0.$$

Proof. Because the difference of two step processes is again a step process, this statement follows from Lemma 3.8. □

Theorem 3.10. *Suppose that* (H_n) *is a sequence of progressively measurable step processes and* H *is a progressively measurable process such that*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_0^\infty (H_n(s) - H(s))^2 ds\right] = 0.$$

Then

$$\int_0^\infty H(s) dB_s := \lim_{n\to\infty} \int_0^\infty H_n(s) dB_s$$

exists as a limit in the L^2 *sense and is independent of the choice of the approximating sequence. Furthermore, we have the Itô isometry*

$$\mathbf{E}\left[\left(\int_0^\infty H(s) dB_s\right)^2\right] = \mathbf{E}\left[\int_0^\infty H(s)^2 ds\right].$$

Proof. Let (H_n) be such a sequence. By the triangle inequality, we have $\mathbf{E}[\int (H_n - H_m)^2 ds]^{1/2} \leq \mathbf{E}[\int (H_n - H)^2 ds]^{1/2} + \mathbf{E}[\int (H_m - H)^2 ds]^{1/2}$ so that we may apply Lemma 3.9 to conclude that $(\int_0^\infty H_n(s) dB_s)$ is a Cauchy sequence. By the completeness of the L^2 space, a limit exists. Lemma 3.9 implies the limit does not depend on the choice of approximating sequence. To see this, assume that (H_n) and (H'_n) are sequences such that $\|H_n - H\|_2 \to 0$ and $\|H'_n - H\|_2 \to 0$, respectively, and that $\lim_{n\to\infty} \int H_n(s) dB_s =: X$ in the L^2 sense. We again have $\mathbf{E}[\int (H_n - H'_n)^2 ds]^{1/2} \leq \mathbf{E}[\int (H_n - H)^2 ds]^{1/2} + \mathbf{E}[\int (H'_n - H)^2 ds]^{1/2}$. Again, since the difference of two step functions is a step function, we have that, by the same logic as Lemma 3.9, $\int (H_n(s) - H'_n(s)) dB_s \to 0$ in L^2. By the triangle inequality in the L^2 norm, we have $\mathbf{E}[(\int H'_n dB_s - X)^2]^{1/2} \leq \mathbf{E}[(\int H'_n dB_s - \int H_n dB_s)^2]^{1/2} + \mathbf{E}[(\int H_n dB_s - X)^2]^{1/2}$. Thus, we see $\lim_{n\to\infty} \int H'_n dB_s = X$ in the L^2 sense, so indeed X is the unique limit $\int H dB_s$. The Itô isometry follows from Lemma 3.8 and taking $n \to \infty$. □

We have therefore proven the Itô integral exists for any progressively measurable process H in the sense stated in Definition 3.6. There are two more desired properties that we would like our integral to have. Like with Riemann integrals, we would like $t \mapsto \int_0^t H_s dB_s$ to be continuous. Furthermore, we would also expect the integral to be a martingale; that is, intuitively, we should not be able to make a profit placing bets on the Brownian motion path. To state the martingale property, we need a notion of a stochastic integral with finite upper limit.

Definition 3.11. Suppose $(H_s)_{s\geq 0}$ is progressively measurable with $\mathbf{E}[\int_0^t H_s^2 ds] < \infty$. We define H^t such that $H^t(s) := H_s \mathbf{1}_{s\leq t}$, and

$$\int_0^t H_s dB_s := \int_0^\infty H^t(s) dB_s.$$

In addition, the martingale property also requires the notion of a modification of a process, which we define now.

Definition 3.12. A ***modification*** of a process $(X_t)_{t\geq 0}$ is a process $(Y_t)_{t\geq 0}$ such that for all $t \geq 0$, $\mathbf{P}(X_t = Y_t) = 1$.

This is not to be confused with two processes agreeing everywhere almost surely; a counterexample is $([0,1], \mathcal{B}, \lambda)$ where λ is Lebesgue measure and $X_t(\omega) = \omega$ if $t = \omega$ and 0 otherwise as a modification of 0. The following result from [2, Theorem 7.11] gives a continuous modification of the stochastic integral.

Theorem 3.13. *Suppose $(H_s)_{s\geq 0}$ is progressively measurable and for any $t \geq 0$ we have $\mathbf{E}[\int_0^t H_s^2 ds] < \infty$. Then there exists a modification of the process $(\int_0^t H_s dB_s)_{t\geq 0}$ that is almost surely a continuous martingale.*

Thus, we have constructed a stochastic integral as defined by Definition 3.6 which has a continuous martingale modification. Therefore, we can identify the process with that modification, so that we regard the stochastic integral as a continuous martingale–an intuitive and desirable property. For instance, the martingale property is used in in the proof of Theorem 4.3 and continuity is used in the proof of Theorem 3.17.

Far more general Itô integrals have been constructed, as is done in texts like [1] and [4], which construct the integral for semimartingales, the largest possible class of processes for which we can define an Itô integral. We instead have mostly followed the construction in [2], which constructs the Itô integral for Brownian motion only. Their construction allows us to obtain many important results with less effort.

3.2. **Itô's Formula.** There are many different versions of Itô's formula. The one we will prove here, given as an exercise in [2, Theorem 7.15], will be perhaps the most general version for Brownian motion, so that we can get all lesser versions simultaneously. Though the proof in full detail is long, it centers around a rather simple idea of performing a Taylor expansion of f, then using the continuity of the derivatives to obtain a bound on the remainder given by the Taylor expansion, then using the convergence of the Riemann sums to their respective integral limits. To get a sense of this general idea, we sketch the proof of the univariate case, outlining the proof in [2, Theorem 7.13]. Much of the effort in the general case is devoted to adapting this general strategy to high dimensions and ensuring all the details still align to constitute a full proof. To provide such a proof, we will need the following convergence result from [2, Theorem 7.12].

Lemma 3.14. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is a continuous function, $t > 0$, $(\mathcal{P}_n)$ is a sequence of partitions of the form $0 = t_1^n < \cdots < t_{k_n}^n = t$ with mesh size $\Delta(n) \to 0$. Then, in probability,*

$$\sum_{i=1}^{k_n - 1} f(B_{t_i^n})(B_{t_{i+1}^n} - B_{t_i^n})^2 \to \int_0^t f(B_s) ds.$$

Theorem 3.15 (Itô's Formula I)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ have a continuous second derivative, and assume that for some $t > 0$ we have $\mathbf{E}[\int_0^t f'(B_s)^2 ds] < \infty$. Then, almost surely, for all $0 \leq s \leq t$, we have*

$$f(B_s) - f(B_0) = \int_0^s f'(B_s) dB_s + \frac{1}{2} \int_0^s f''(B_s) ds.$$

Proof Sketch. Defining

$$\omega(\delta, M) := \sup_{\substack{x_1, x_2 \in [-M, M] \\ |x_1 - x_2| < \delta}} |f''(x) - f''(y)|,$$

Taylor's theorem implies that for all y close enough to x,

$$|f(y) - f(x) - f'(x)(y-x) - \frac{1}{2} f''(x)(y-x)^2| \leq \omega(\delta, M)(y-x)^2.$$

Then with appropriate choice of δ and M, for any partition,

$$\begin{aligned} |f(B_t) - f(B_0) - &\sum_{i=1}^{n-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) \\ &- \frac{1}{2} \sum_{i=1}^{n-1} f''(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2| \leq \omega(\delta, M) \sum_{i=1}^{n-1} (B_{t_{i+1}} - B_{t_i})^2. \end{aligned}$$

Then, take $n \to \infty$. On the LHS, one gets convergence of these sums to their respective integrals while the RHS converges to 0, giving the statement for $s = t$. Repeat for each $s \in \mathbb{Q}$ and exploit continuity of the left and right side of Itô's formula as a function of s to get the statement on $[0, t]$. □

To state and prove the multivariable Itô's formula, it is helpful to introduce the following notation.

Notation 3.16. We consider $f : \mathbb{R}^{d+m} \to \mathbb{R}$ as a function $f(x, y)$ where $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^m$. We have $\nabla_x f = (\partial_1 f, ..., \partial_d f)$ and $\nabla_y f = (\partial_{d+1} f, ..., \partial_{d+m} f)$. We write

$$\int_0^t \nabla_x f(B_u, \zeta_u) \cdot dB_u = \sum_{i=1}^d \int_0^t \partial_i f(B_u, \zeta_u) B_u^i,$$
$$\int_0^t \nabla_y f(B_u, \zeta_u) \cdot d\zeta_u = \sum_{i=1}^m \int_0^t \partial_{d+i} f(B_u, \zeta_u) \zeta_u^i,$$

and

$$\Delta_x f = \sum_{i=1}^d \partial_{ii} f.$$

Lastly, we use the shorthand $a \wedge b$ to denote the minimum of a and b, and we use $a \vee b$ to denote the maximum.

Theorem 3.17 (Multivariable Itô's Formula)**.** *Let (B_t) be a d-dimensional Brownian motion and suppose (ζ_s) is a continuous, adapted, m-dimensional stochastic process whose components are all increasing. Let $f : \mathbb{R}^{d+m} \to \mathbb{R}$ be a function such that the partial derivatives $\partial_i f$ and $\partial_{jk} f$ exist for all $1 \leq j, k \leq d$ and $d + 1 \leq i \leq d + m$ and are continuous. If, for some $t > 0$,*

$$\mathbf{E}\left[\int_0^t |\nabla_x f(B_s, \zeta_s)|^2 ds\right] < \infty,$$

then, almost surely, for all $0 \le s \le t$,

$$f(B_s,\zeta_s) - f(B_0,\zeta_0) = \int_0^s \nabla_x f(B_u,\zeta_u)\cdot dB_u + \int_0^s \nabla_y f(B_u,\zeta_u)\cdot d\zeta_u + \frac{1}{2}\int_0^s \Delta_x f(B_u,\zeta_u)du. \tag{3.18}$$

Proof. First, note we have that Lemma 3.14 also holds when f has a d-dimensional Brownian motion and an m-dimensional adapted process in its argument. That is, we have for a sequence of partitions $0 = t_1^n < \cdots < t_{k_n}^n = t$ of $[0,t]$ with mesh size going to zero, in probability,

$$\lim_{n\to\infty}\sum_{i=1}^{k_n-1} f(B_{t_i^n},\zeta_{t_i^n})(B^j_{t_{i+1}^n} - B^j_{t_i^n})^2 = \int_0^t f(B_s,\zeta_s)ds,$$

where B_s^j is the jth component of B_s. Now, let $\mathrm{Hes}_x f(x,y)$ denote the Hessian matrix $\left[\frac{\partial^2}{\partial x_i \partial x_j} f(x,y)\right]_{ij}$ and define

$$\omega_1(\delta,M) := \sup_{\substack{x_1,x_2\in[-M,M]^d\\ y_1,y_2\in[-M,M]^m\\ |x_1-x_2|\vee|y_1-y_2|<\delta}} |\nabla_y f(x_1,y_1) - \nabla_y f(x_2,y_2)|,$$

$$\omega_2(\delta,M) := \sup_{\substack{x_1,x_2\in[-M,M]^d\\ y_1,y_2\in[-M,M]^m\\ |x_1-x_2|\vee|y_1-y_2|<\delta}} \|\mathrm{Hes}_x f(x_1,y_1) - \mathrm{Hes}_x f(x_2,y_2)\|,$$

where $\|\cdot\|$ denotes the operator norm. Now take $x, x_0 \in [-M,M]^d$ and $y, y_0 \in [-M,M]^m$ with $|x-x_0|\vee|y-y_0| < \delta$. By the multivariate mean value theorem, there exists $\tilde{y}$ on the line segment connecting y and y_0 with the property that $|\tilde{y}-y|\vee|\tilde{y}-y_0| < \delta$ such that

$$f(x,y) - f(x,y_0) = \nabla_y f(x,\tilde{y})\cdot(y-y_0).$$

Hence,

$$|f(x,y) - f(x,y_0) - \nabla_y f(x_0,y_0)\cdot(y-y_0)| \le \omega_1(\delta,M)|y-y_0|.$$

Similarly, by the multivariate Taylor's theorem, we have that

$$f(x,y_0) = f(x_0,y_0) + \nabla_x f(x_0,y_0)(x-x_0) + \frac{1}{2}(x-x_0)^T\mathrm{Hes}_x f(\tilde{x},y_0)(x-x_0)$$

for some $\tilde{x}$ on the line segment connecting x and x_0. Therefore,

$$\begin{aligned}
&|f(x,y_0) - f(x_0,y_0) - \nabla_x f(x_0,y_0)\cdot(x-x_0) - \frac{1}{2}(x-x_0)^T\mathrm{Hes}_x f(x_0,y_0)(x-x_0)|\\
&= |\frac{1}{2}(x-x_0)^T\mathrm{Hes}_x f(x_0,y_0)(x-x_0) - \frac{1}{2}(x-x_0)^T\mathrm{Hes}_x f(\tilde{x},y_0)(x-x_0)|\\
&\le |(x-x_0)^T(\mathrm{Hes}_x f(x_0,y_0) - Hes_x f(\tilde{x},y_0))(x-x_0)|\\
&\le \|\mathrm{Hes}_x f(x_0,y_0) - \mathrm{Hes}_x f(\tilde{x},y_0)\||x-x_0|^2\\
&\le \omega_2(\delta,M)|x-x_0|^2,
\end{aligned}$$

where the fourth line follows from the third line by observing that, from the Cauchy-Schwarz inequality and from the fact that $|Ax| \le \|A\||x|$, we have $|x^TAx| \le |x^T||Ax| \le \|A\||x|^2$, where A is a matrix and x is a vector of appropriate dimension.

It then follows from the triangle inequality that

$$\begin{aligned}|f(x,y) - f(x_0,y_0) - \nabla_y f(x_0,y_0)\cdot(y-y_0) - \nabla_x f(x_0,y_0)\cdot(x-x_0)\\ - \frac{1}{2}(x-x_0)^T \mathrm{Hes}_x f(x_0,y_0)(x-x_0)|\\ \leq \omega_1(\delta,M)|y-y_0| + \omega_2(\delta,M)|x-x_0|^2.\end{aligned}$$

Now, for any partition $0 = t_1 < t_2 < \cdots < t_n = t$, define

$$\delta := \max_{1\leq i\leq n-1}|B_{t_{i+1}} - B_{t_i}| \wedge \max_{1\leq i\leq n-1}|\zeta_{t_{i+1}} - \zeta_{t_i}|,$$

$$M := \max_{0\leq s\leq t}|B_s| \wedge \max_{0\leq s\leq t}|\zeta_s|.$$

In the above inequality, replace x with $B_{t_{i+1}}$, x_0 with B_{t_i}, y with $\zeta_{t_{i+1}}$, and y_0 with ζ_{t_i}, and sum all of the $n-1$ terms of the partition to conclude that

$$\begin{aligned}&|f(B_t,\zeta_t) - f(B_0,\zeta_0) - \sum_{i=1}^{n-1}\nabla_x f(B_{t_i},\zeta_{t_i})\cdot(B_{t_{i+1}} - B_{t_i})\\ &- \sum_{i=1}^{n-1}\nabla_y f(B_{t_i},\zeta_{t_i})\cdot(\zeta_{t_{i+1}} - \zeta_{t_i}) - \frac{1}{2}\sum_{i=1}^{n-1}(B_{t_{i+1}} - B_{t_i})^T \mathrm{Hes}_x f(B_{t_i},\zeta_{t_i})(B_{t_{i+1}} - B_{t_i})|\\ &\leq \omega_1(\delta,M)\|\zeta_t - \zeta_0\|_1 + \omega_2(\delta,M)\sum_{i=1}^{n-1}|B_{t_{i+1}} - B_{t_i}|^2.\end{aligned}$$

To see why the first term of the upper bound holds, we observe that

$$\begin{aligned}\sum_{i=1}^{n-1}|\zeta_{t_{i+1}} - \zeta_{t_i}| &\leq \sum_{i=1}^{n-1}\sum_{j=1}^{m}|\zeta^j_{t_{i+1}} - \zeta^j_{t_i}|\\ &= \sum_{j=1}^{m}\sum_{i=1}^{n-1}(\zeta^j_{t_{i+1}} - \zeta^j_{t_i})\\ &= \sum_{j=1}^{m}(\zeta^j_t - \zeta^j_0)\\ &= \sum_{j=1}^{m}|\zeta^j_t - \zeta^j_0|.\end{aligned}$$

Now, consider the following sums:

$$\sum_{i=1}^{n-1}\nabla_x f(B_{t_i},\zeta_{t_i})\cdot(B_{t_{i+1}} - B_{t_i}) = \sum_{k=1}^{d}\sum_{i=1}^{n-1}\partial_k f(B_{t_i},\zeta_{t_i})(B^k_{t_{i+1}} - B^k_{t_i}), \tag{3.19}$$

$$\sum_{i=1}^{n-1}\nabla_y f(B_{t_i},\zeta_{t_i})\cdot(\zeta_{t_{i+1}} - \zeta_{t_i}) = \sum_{k=1}^{m}\sum_{i=1}^{n-1}\partial_{d+k} f(B_{t_i},\zeta_{t_i})(\zeta^k_{t_{i+1}} - \zeta^k_{t_i}), \tag{3.20}$$

$$\sum_{i=1}^{n-1}(B_{t_{i+1}}-B_{t_i})^T \mathrm{Hes}_x f(B_{t_i},\zeta_{t_i})(B_{t_{i+1}}-B_{t_i})$$

$$(3.21)\qquad = \sum_{j,k=1}^{d}\sum_{i=1}^{n-1}\frac{\partial^2}{\partial x_j \partial x_k} f(B_{t_i},\zeta_{t_i})(B^j_{t_{i+1}}-B^j_{t_i})(B^k_{t_{i+1}}-B^k_{t_i}).$$

By definition of the stochastic integral, as the mesh size goes to zero, (3.19) converges in the L^2 sense to $\sum_{k=1}^d \int_0^t \partial_k f(B_u,\zeta_u)dB_u = \int_0^t \nabla_x f(B_u,\zeta_u)dB_u$. Next, (3.20) converges almost surely to $\int_0^t \nabla_y f(B_u,\zeta_u)d\zeta_u$ by Theorem 3.1. Considering (3.21), we also claim that $(B^j_{t_{i+1}}-B^j_{t_i})(B^k_{t_{i+1}}-B^k_{t_i})$ converges to 0 in probability when $j \neq k$. Since these Brownian motions are independent, we have

$$\mathbf{E}[(B^j_{t_{i+1}}-B^j_{t_i})(B^k_{t_{i+1}}-B^k_{t_i})] = \mathbf{E}[(B^j_{t_{i+1}}-B^j_{t_i})]\mathbf{E}[(B^k_{t_{i+1}}-B^k_{t_i})] = 0.$$

Furthermore, since the independent increments are each of mean 0, we have

$$\begin{aligned}\mathrm{Var}[(B^j_{t_{i+1}}-B^j_{t_i})(B^k_{t_{i+1}}-B^k_{t_i})] &= \mathrm{Var}[(B^j_{t_{i+1}}-B^j_{t_i})]\mathrm{Var}[(B^k_{t_{i+1}}-B^k_{t_i})]\\ &= (t_{i+1}-t_i)^2\\ &\leq \Delta(n)^2\\ &\to 0,\end{aligned}$$

proving L^2 convergence, hence convergence in probability. It therefore follows from the above and from the multivariable version of Lemma 3.14 that (3.21) converges in probability to $\sum_{k=1}^d \int_0^t \frac{\partial^2}{\partial x_k^2} f(B_u,\zeta_u)du = \int_0^t \Delta_x f(B_u,\zeta_u)du$.

Next, we observe that $\sum_{i=1}^{n-1}|B_{t_{i+1}}-B_{t_i}|^2 = \sum_{j=1}^d \sum_{i=1}^{n-1}(B^j_{t_{i+1}}-B^j_{t_i})^2$ (switching the order of summation), which converges to dt in probability by Theorem 2.7. To get almost sure convergence for all terms as $n \to \infty$, we observe the following: First, take a nested sequence of partitions. Then, Theorem 2.7 in fact gives almost sure convergence to dt. Next, by continuity of Brownian motion and of ζ, ω_1 and ω_2 converge to 0 almost surely. Second, convergence in probability of (3.19)+(3.20)+(3.21) holds for this nested sequence of partitions in particular, and we recall that there exists a subsequence of partitions such that convergence of these terms to the sum of their integral limits holds almost surely. This subsequence is also a sequence of nested partitions, so almost sure convergence of $\sum_{i=1}^{n-1}|B_{t_{i+1}}-B_{t_i}|^2$ still holds. Itô's formula therefore holds for $s = t$ almost surely.

We now prove the formula for $0 \leq s < t$. By monotonicity of the integral, if $s < t$, and s is rational, then $\mathbf{E}[\int_0^s |\nabla_x f(B_s,\zeta_s)|^2 ds] \leq \mathbf{E}[\int_0^t |\nabla_x f(B_s,\zeta_s)|^2 ds] < \infty$. Then we can apply the above argument to get (3.18) with upper limit of s. Since s was arbitrary, this holds for each rational s. Since the intersection of countably many almost sure events is almost sure, we therefore get that (3.18) almost surely holds for all rational s between 0 and t simultaneuously. Both sides of (3.18) are continuous functions of s by Theorem 3.13, hence their equality on rationals in $[0,t]$ implies their equality on the whole interval, proving the statement. □

Remark 3.22. Often, Itô's formula will be stated in terms of differentials. For instance, the one-dimensional statement may be written as an equation

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt.$$

Such equations themselves don't have precise mathematical meaning, but they are in reference to integral equations, such as those seen in Theorem 3.15, which do. Texts will often use the differential versions to do computations, with the understanding that such expressions are referring to the integral forms.

4. Applications

In general, exact computations of stochastic integrals are not feasible. However, Itô's formula allows us to compute a simple example.

Example 4.1. What is the integral of Brownian Motion with respect to itself? Let's consider the single variable version of Itô's formula and let $f(x) = x^2$ on $[0, t]$, so that the derivative will just be a constant multiple of the identity function. Then Itô's formula gives

$$\begin{aligned} B_t^2 &= \int_0^t f'(B_u)dB_u + \frac{1}{2}\int_0^t f''(B_u)du \\ &= 2\int_0^t B_u dB_u + \int_0^t du, \end{aligned}$$

hence $\int_0^t B_u dB_u = B_t^2/2 - t/2$.

We now discuss how Brownian motion relates to the heat equation via the Feynman-Kac formula. To begin, we define the conditions of the heat equation.

Definition 4.2. Let $U \subset \mathbb{R}^d$ be open and bounded. We say that $u : [0, \infty) \times U \to [0, \infty)$ ***solves the heat equation with initial condition*** $g : U \to [0, \infty)$ on U if we have

$$\lim_{\substack{x \to x_0 \\ t \to 0^+}} u(t, x) = f(x_0) \text{ if } x_0 \in U,$$

$$\lim_{\substack{x \to x_0 \\ t \to t_0}} u(t, x) = 0, \text{ if } x_0 \in \partial U,$$

$$\partial_t u(t, x) = \frac{1}{2}\Delta_x u(t, x) \text{ on } (0, \infty) \times U.$$

This describes a scenario where particles have initial temperature given by f, particles are killed, i.e., lose all heat, at the boundary, and diffuse over time according to a rule where a particle at x compares itself to its neighbors in the sense that, if the neighbors in an infinitesimal ball around x on average have more heat than x, then heat flows to x. If the opposite holds, then heat flows away from x. One can discuss solutions to such equations in terms of Brownian motion with Feynman-Kac formulas, of which there are several. The following simplified Feynman-Kac formula is stated and proven in [2].

Theorem 4.3 (Feynman-Kac)**.** *Suppose u is a bounded, twice continuously differentiable solution of the heat equation on the domain U with continuous initial condition g. Then*

$$u(t, x) = \mathbf{E}_x[g(B_t)\mathbf{1}\{t < \tau\}],$$

where τ is the exit time of the Brownian motion from U.

Proof. One can also apply Theorem 3.17 to stopping times of the form $s \wedge T$ where $s \leq t$, and T is the exit time from a compact set $K \subset U$ (for a discussion, see [2, pp. 200]). Now, consider a compact $K \subset U$ and let σ be the respective escape time when starting from x. Fix $t > 0$. We apply (3.18) with $f(x, y) = u(t - y, x)$ and $\zeta_s = s$. We have, for any $s < t$,

$$u(t - (s \wedge \sigma), B_{s\wedge\sigma}) - u(t, B_0) = \int_0^{s\wedge\sigma} \nabla_x u(t - v, B_v) \cdot dB_v - \int_0^{s\wedge\sigma} \partial_t u(t - v, B_v) dv + \frac{1}{2} \int_0^{s\wedge\sigma} \Delta_x u(t - v, B_v) dv,$$

with the chain rule being applied to the second integral. However, by the third property of the heat equation, the second and third integrals cancel. Take the expectation of both sides; the expectation of the stochastic integral is 0, even when the upper limit of integration is a stopping time, hence

$$\mathbf{E}_x[u(t - (s \wedge \sigma), B_{s\wedge\sigma})] = \mathbf{E}_x[u(t, B_0)] = u(t, B_0) = u(t, x).$$

Applying the law of total expectation by distinguishing $\{s < \sigma\}$ and $\{s \geq \sigma\}$, and considering as σ increases to τ (by having K increase to U), we obtain $\mathbf{E}_x[u(t - s, B_s)\mathbf{1}\{s < \tau\}] = u(t, x)$. Now, take a limit $s \to t^-$ to obtain the statement. □

Example 4.4. Take $g = 1$ on U, representing a start of unit temperature. Then $u(t, x) = \mathbf{P}_x[t < \tau]$. The temperature at x at time t can be thought of as the proportion of Brownian motion paths that started at x and have not yet been killed off at the boundary at time t.

Our next application of Brownian motion is the conformal invariance property. A difficult but self-contained proof is given in [2], which uses Itô's formula. To provide a different proof, we use the Dubins-Schwarz theorem, which is proven in [6, p. 181] and relies on the notion of a local martingale, which we introduce now.

Definition 4.5. An adapted process $(X_t)_{0\leq t\leq T}$ is a ***local martingale*** if there exists a sequence of stopping times (T_n) that is almost surely increasing to T such that the process $(X_{t\wedge T_n})_{t\geq 0}$ is a martingale for every n.

Theorem 4.6 (Dubins-Schwarz)**.** *Let (M_s) be a continuous local martingale for which $M_0 = 0$ almost surely and $\lim_{t\to\infty}[M]_t = \infty$ almost surely. Let $\sigma(t) := \inf\{s \geq 0 : [M]_s > t\}$. Then for all $t \geq 0$, $\sigma(t)$ is an $(\mathcal{F}(s))_{s\geq 0}$ stopping time. Furthermore, $(\mathcal{F}(\sigma(t)))_{t\geq 0}$ is a filtration, and $M_{\sigma(t)}$ is a Brownian motion adapted to $(\mathcal{F}(\sigma(t)))_{t\geq 0}$.*

The Dubins-Schwarz Theorem is telling us that if we are able to see how the quadratic variation of the process is changing over time, we are able to "reverse-engineer" a Brownian motion by reparameterizing the process according to that variation. For example, if $[M]_t = t^2$, then we would have that $[M]_s > t$ as soon as $s > \sqrt{t}$, that is $\sigma(t) = \sqrt{t}$, and $M_{\sqrt{t}}$ would be a Brownian motion.

We now state a sufficient condition for obtaining a local martingale. This will be useful for when we encounter holomorphic functions, which have harmonic real and imaginary parts.

Proposition 4.7. *Let $D \subset \mathbb{R}^d$ be a domain and $f : D \to \mathbb{R}$ be harmonic on D. Suppose that $(B_t)_{0\leq t\leq T}$ is a Brownian motion started in D and stopped at* *the time T when the Brownian motion first exits D. Then $(f(B_t))_{0\le t\le T}$ is a local martingale.*

We would also like to have a sufficient condition for when a local martingale is actually a martingale. The next result tells us that it is enough to have the local martingale be bounded (see [4, pp. 37-38]).

Proposition 4.8. *If (X_t) is a bounded local martingale, then (X_t) is in fact a martingale.*

Before proving conformal invariance, we prove a general fact about covariation of Brownian motions. Recall from Definition 2.6 that covariation is defined as a limit in probability.

Lemma 4.9. *Let (X_t) and (Y_t) be independent linear Brownian motions. Then $[X,Y]_t = 0$.*

Proof. It is enough to show that the L^2 limit of the sum in Definition 2.6 is 0. Consider a sequence of partitions $(\mathcal{P}_n)$ of the form $0 = t_1^n < \cdots < t_{k_n}^n = t$ of $[0,t]$ and denote $W_i := X_{t_{i+1}^n} - X_{t_i^n}$ and $Z_i := Y_{t_{i+1}^n} - Y_{t_i^n}$. Then

$$\mathbf{E}\left[\sum_{i=1}^{k_n-1} W_i Z_i\right] = \sum_{i=1}^{k_n-1} \mathbf{E}[W_i]\mathbf{E}[Z_i] = 0,$$

and

$$\begin{aligned}\mathbf{E}\left[\left(\sum_{i=1}^{k_n-1} W_i Z_i\right)^2\right] &= \sum_{i=1}^{k_n-1} \mathbf{E}[W_i^2 Z_i^2] + 2\sum_{i<j} \mathbf{E}[W_i Z_i W_j Z_j] \\ &= \sum_{i=1}^{k_n-1} (t_{i+1}^n - t_i^n)^2 + 2\sum_{i<j} \mathbf{E}[W_i]\mathbf{E}[Z_i]\mathbf{E}[W_j]\mathbf{E}[Z_j] \\ &\le \Delta(n)t \\ &\to 0.\end{aligned}$$

Thus the sum converges to 0 in L^2 and hence in probability. □

Theorem 4.10 (Conformal Invariance)**.** *Let $D \subset \mathbb{C}$ be a domain and let (B_t) be a Brownian motion started at $z \in \mathbb{C}$. If $f : D \to \mathbb{C}$ is holomorphic, then there exists a Brownian motion $\tilde{B}_t$ in $f(D)$ and started in $f(z)$ such that $f(B_t) = \tilde{B}_{\int_0^t |f'(B_s)|^2 ds}$.*

Proof of Theorem. The main task is to compute the quadratic variation of $f(B_t)$, which is done with Itô's formula, and the covariation of a vector is determined pairwise, in this case $[u,u], [v,v]$, and $[u,v]$. We split everything into real and imaginary parts. Writing $z = x + iy$, we have $f(z) = u(x,y) + iv(x,y)$, and $B_t = X_t + iY_t$, where X_t and Y_t are independent linear Brownian motions. We recall the Cauchy-Riemann equations: $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$ and $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$, and the fact that both u and v are harmonic. The multivariable Itô's formula Theorem 3.17 implies that

$$\begin{aligned} du(X_t,Y_t) &= \frac{\partial}{\partial x}u(X_t,Y_t)dX_t + \frac{\partial}{\partial y}u(X_t,Y_t)dY_t + \frac{1}{2}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})u(X_t,Y_t)dt \\ &= \frac{\partial}{\partial x}u(X_t,Y_t)dX_t + \frac{\partial}{\partial y}u(X_t,Y_t)dY_t, \qquad (4.11)\end{aligned}$$

by harmonicity. In exactly the same way, one has

$$dv(X_t, Y_t) = \frac{\partial}{\partial x} v(X_t, Y_t) dX_t + \frac{\partial}{\partial y} v(X_t, Y_t) dY_t.$$

We have, by using (4.11) and considering the differential of the quadratic variation process $[u(B_t)]$,

$$\begin{aligned} d[u(B_t)] &= \frac{\partial}{\partial x} u(B_t)^2 d[X_t, X_t] + 2\frac{\partial u}{\partial x} u(B_t) \frac{\partial}{\partial y}(B_t) d[X_t, Y_t] + \frac{\partial}{\partial y} u(B_t)^2 d[Y_t, Y_t] \\ &= \frac{\partial}{\partial x} u(B_t)^2 dt + \frac{\partial}{\partial y} u(B_t)^2 dt \\ &= |f'(B_t)|^2, \end{aligned}$$

and by identical calculations, one obtains

$$\begin{aligned} d[v(B_t)] &= \frac{\partial}{\partial x} v(B_t)^2 dt + \frac{\partial}{\partial y} v(B_t)^2 dt \\ &= \frac{\partial}{\partial y} u(B_t)^2 dt + \frac{\partial}{\partial x} u(B_t)^2 dt \\ &= d[u(B_t)], \end{aligned}$$

where the second equality follows from the first by applying the Cauchy-Riemann equations. We conclude that $[u(B_s)]_t = [v(B_s)]_t = \int_0^t |f'(B_s)|^2 ds$, and $[u(B_s), v(B_s)]_t = 0$.

Now, let $\sigma(t) := \inf\{s \geq 0 : \int_0^s |f'(B_u)|^2 du > t\}$, and define $\tilde{B}_t = f(B_{\sigma(t)}) = u(B_{\sigma(t)}) + iv(B_{\sigma(t)})$. By Theorem 4.6, $\tilde{B}_t$ is a Brownian motion with respect to $\mathcal{F}(\sigma(t))$. This implies the statement, because $\tilde{B}_{\int_0^t |f'(B_s)|^2 ds} = f(B_{\sigma(\int_0^t |f'(B_s)|^2 ds)})$, but it is clear from the increasing of the integral and the definition of σ that $\sigma(\int_0^t |f'(B_s)|^2 ds) = t$. (The isolated nature of the zeros of f' implies that B_s will not stay in a region where $f' = 0$ for any time interval, so the integral is indeed increasing). □

To see more ways that Itô's formula and its corollaries relate to complex analysis, we now offer an alternative proof of Liouville's theorem, following [7, Theorem 2.4.5].

Theorem 4.12 (Liouville's Theorem)**.** *Assume that $f : \mathbb{C} \to \mathbb{C}$ is a holomorphic bounded function. Then f is constant.*

Proof. We use conformal invariance. Our assumption implies that for a Brownian motion (B_t) in the complex plane, $t \mapsto f(B_t) = \tilde{B}_{\int_0^t |f'(B_s)|^2 ds}$ is a bounded function, but since Brownian motion visits every neighborhood by Theorem 2.16, this must mean that $t \mapsto \int_0^t |f'(B_s)|^2 ds$ is bounded. If f is constant, we are done, so assume that f is nonconstant. We can then consider a disk D whose closure contains no zero of f'. Furthermore, there exists $\delta > 0$ such that $|f'(z)| \geq \delta$ for each $z \in D$. Define S_n and T_n to respectively be the nth entrance and exit times of $\bar{D}$. The Strong Markov property implies that $T_n - S_n$ are i.i.d. random variables. Furthermore, they are of finite expectation; take a rectangle $[a, b] \times [c, d]$ containing D. The exit time from the rectangle is less than or equal to the time the first component exits $[a, b]$, and Proposition 2.15 tells us the expectation of the latter is finite. We have $\int_0^\infty |f'(B_s)|^2 ds \geq \sum_{n=1}^\infty \delta^2 (T_n - S_n) = \infty$ almost surely, where the last equality follows from the Strong Law of Large Numbers. □

We conclude with two proofs of the Fundamental Theorem of Algebra, again following [7, Theorem 2.4.3]. While the second proof provided below does not use Itô's formula, it is a slick argument that further illustrates the utility of Brownian motion in complex analysis.

Theorem 4.13 (The Fundamental Theorem of Algebra)**.** *Assume that* $p : \mathbb{C} \to \mathbb{C}$ *is a nonconstant polynomial. Then there exists* z_0 *such that* $p(z_0) = 0$.

Proof. Assume the contrary. Define $f(z) = 1/p(z)$ on all of $\mathbb{C}$. Then f is holomorphic, and since $|p(z)| \to \infty$ as $|z| \to \infty$, it then follows that f is bounded on $\mathbb{C}$. It is at this point that we can use Theorem 4.12 and be done: f would then be constant, a contradiction, as goes the usual proof of this theorem. But we can also prove the theorem using facts about Brownian motion.

It follows from Proposition 4.7 and Proposition 4.8 that $M_t := \Re f(B_t)$ is a bounded local martingale and is therefore a martingale (we use $\Re f$ to denote the real part of f). Since it is bounded, we can also apply the Martingale Convergence Theorem to conclude that there exists some M_∞ such that $\lim_{t\to\infty} M_t = M_\infty$ almost surely. Since $(M_t)_{t\geq 0}$ is bounded and nonconstant (using that Brownian motion visits every neighborhood, and $\Re f(z)$ being nonconstant follows from the Cauchy-Riemann equations), consider any a, b such that $\inf M_t < a < b < \sup M_t$. Consider the disjoint sets $E_1 := \{t : M_t < a\}$ and $E_1 := \{t : M_t > b\}$. Since $(B_t)_{t\geq 0}$ visits every neighborhood of the plane, these sets are nonempty, and by continuity of the process (M_t), they are open. Lastly, (B_t) visits each neighborhood infinitely often, in particular the neighborhoods contained in $\Re f^{-1}(-\infty, a)$ and $\Re f^{-1}(b, \infty)$. Therefore, we conclude that

$$\liminf_{t\to\infty} M_t \leq a < b \leq \limsup_{t\to\infty} M_t,$$

contradicting convergence and proving the statement. ☐

Acknowledgements

It is a pleasure to thank my mentor, Manan Bhatia, for his support and guidance throughout this program, especially with helping me gain intuition for much of the content covered here. Lastly, thank you to Peter May for running the REU program that makes papers like this possible.

References

[1] Durrett, Richard. Stochastic Calculus: A Practical Introduction. CRC Press, 1996.

[2] Mörters, Peter, and Y. Peres. Brownian Motion. Cambridge University Press, 2012, https://people.bath.ac.uk/maspm/book.pdf, Accessed 29 Aug. 2023.

[3] Protter, Murray H., and Charles B. Morrey. A First Course in Real Analysis. Springer, 1991.

[4] Protter, Philip. Stochastic Integration and Differential Equations. Springer, 2005.

[5] Resnick, Sidney. Adventures in Stochastic Processes. Birkhäuser, 2005.

[6] Revuz, D., and M. Yor. Continuous Martingales and Brownian Motion. Springer, 1999.

[7] Watson, Sam. "Brownian Motion, Complex Analysis, and the Dimension of the Brownian Frontier." Cambridge University, 2010.