STOCHASTIC DIFFERENTIAL EQUATIONS AND FINANCIAL APPLICATIONS

INANNA KHANSA

Abstract. Stochastic differential equations are essential to modeling the random nature of the world, and thus, we take interest in solving them. This paper will introduce the basics of Brownian motion and Itô Calculus to then prove the existence and uniqueness of solutions to a particular class of stochastic differential equations. With this background, we will then discuss an application to options trading. We assume a familiarity with basic probability theory and real analysis.

1. Introduction

Differential equations are critical to modeling systems that evolve over time, but oftentimes, these deterministic models cannot accurately describe the many factors and variables of a system. This necessitates random or *stochastic* differential equations.

Consider the deterministic ordinary differential equation

$$\begin{cases} \dot{x}(t) = b(x(t), t) & t > 0 \\ x(0) = x_0, \end{cases} \tag{1.1}$$

where $b : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ is a smooth vector field. The solution to this equation would be the smooth trajectory or path $x(\cdot) \colon [0, \infty) \to \mathbb{R}^n$.

Since it is useful to study objects that are not so well-behaved, we define the stochastic analog of the differential equation above as

Date: DEADLINES: Draft AUGUST 14 and Final version AUGUST 28, 2021.

$$\begin{cases} \dot{x}(t) = b(x(t), t) + \sigma(x(t), t)\xi(t) & t > 0 \\ x(0) = x_0. \end{cases}$$

Here, the dynamics of the system are driven by the vector field $b(\cdot)$, and we define $\xi(\cdot)$ to be our stochastic term, which accounts for all that can not be determined. More concretely, we think of $\xi(\cdot)$ as equivalent to $\dot{W}(t)$, or the time derivative of *Brownian motion*, defined in Section 2. Using this intuition, we write our differential equation in differential form[1]:

$$\begin{cases} dx(t) = b(x(t), t)dt + \sigma(x(t), t)dW_t & t > 0 \\ x(0) = x_0. \end{cases} \tag{1.2}$$

In this paper, we will prove that the stochastic differential equation in the form above has a unique solution. We must first, however, rigorously define $\xi(\cdot)$ through Brownian motion and develop the rules of stochastic calculus, which are distinct from their deterministic analog.

2. Brownian Motion

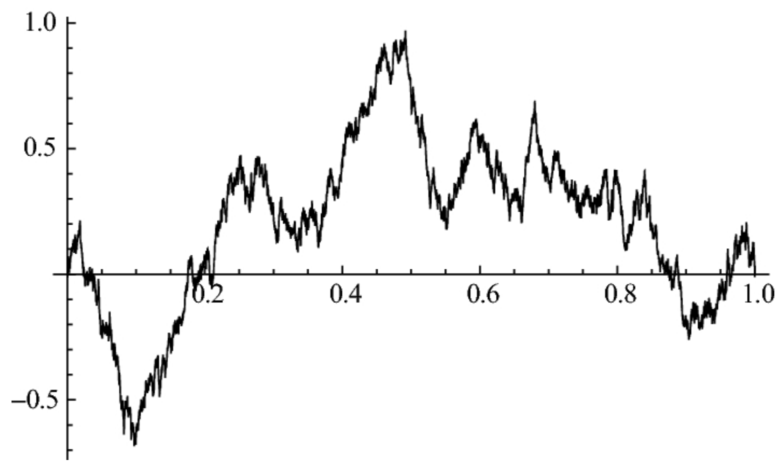

Figure 1. An example of a particular type of Brownian motion, a Brownian bridge.

In essence, "white noise" is the random movement of particle, which can be modeled by a *random walk*. Random walks are a discrete *stochastic processes* that model the position of a particle after taking a certain amount of random steps. When we extend the premise of a discrete random walk to an almost surely continuous stochastic process, we get Brownian motion.

Definition 2.1. A *stochastic process* is a family of random variables indexed by time.

Definition 2.2. A stochastic process $\{W(t) : t \geq 0\}$ is a *(linear) Brownian motion* or *Wiener process* starting at $x \in \mathbb{R}$ if:

[1]Here, dW_t denotes $dW(t)$, and we use this notation in the interest of brevity. In all other cases, this paper will use subscripts to notate indices unrelated to time.

(1) $W(0) = x$. This process is called a *standard Brownian motion* if $W(0) = 0$,
(2) the process has *independent increments*, that is for all times $0 \leq t_1 \leq \cdots \leq t_n$, $W(t_1), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ are independent random variables. Intuitively, this can be thought of as the future motion of a random particle being independent of its motion in the past.
(3) for all times $t \geq s \geq 0$, the increments $W(t) - W(s)$ are normally distributed with mean or expectation zero and variance $t - s$. This property implies that Brownian motion has *stationary increments*, namely that the distribution of an increment relies only on the amount of change in time and is independent of the increment's particular start and end time.
(4) it then follows in particular from properties (2) and (3), that the process has almost surely continuous sample paths or that $t \mapsto W(t)$ is almost surely continuous.

We omit the proof of the existence of such a process, which can be found in [2] on page 46 but include a simple sketch. First, we inductively choose normally distributed random variables $W(t) \sim \mathcal{N}(0, t)$, where t is a dyadic time, with independent increments. Then, we use Gaussian estimates and the Borel-Cantelli lemma[2] to show that the process we constructed is α-Hölder continuous on the dyadic rationals for $\alpha < \frac{1}{2}$. Finally, as we have shown that the process is uniformly continuous, a consequence of Hölder continuity, on a dense set, we uniquely extend the definition of $W(t)$ to all other times t.

With this definition, we move on to helpful and integral properties of Brownian motion.

Definition 2.3. A function $f : [0, \infty) \to \mathbb{R}$ is said to be $\alpha-$*Hölder continuous* on its domain if there exists $\epsilon > 0$ and $c > 0$ such that $|f(x) - f(y)| \leq c|x - y|^\alpha$ for all $x, y \in [0, \infty)$ with $|y - x| < \epsilon$.

Theorem 2.4. *Any sample path of a Brownian motion $t \mapsto W(t, \omega)$ is $\alpha-$Hölder continuous, if $\alpha < \frac{1}{2}$. Namely, for $\alpha < \frac{1}{2}$, there exists a constant $C \in \mathbb{R}$ such that $|W(t) - W(s)| \leq C|t - s|^\alpha$, for all $s, t \in [0, T]$.*

Theorem 2.5. *Brownian motion is nowhere differentiable.*

This property, whose proof follows from Hölder continuity and can be found in [1] on page 18, is why Brownian motion, as in Figure 1, looks so jagged and rough. This theorem also provides the motivation for Itô Calculus as we must define new rules of calculus to make sense of dW_t.

Definition 2.6. A process is called a *time-homogeneous Markov process* if it has the same distribution invariant of its starting point. A Markov process is said to have the *Markov property.*

Intuitively, a stochastic process, $\{X(t) : t \geq 0\}$, has the Markov property if knowing its behavior on an interval $[0, s)$ is not useful at all for predicting its future behavior, i. e. $\{X(t) : t \geq s\}$ as the process "forgets its past." The only factor that affects the process' future path is its present state.

Definition 2.7. A *filtration* on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $\{\mathcal{F}(t) : t \geq 0\}$ of $\sigma-$algebras such that $\mathcal{F}(s) \subset \mathcal{F}(t) \subset \mathcal{F}$, for all $s < t$.

[2]Lemma 5.1

One should think of the filtration generated by Brownian motion until time t as the "information" or the past history of the stochastic process until time t.

Definition 2.8. A random variable τ with values in $[0, \infty)$, defined on a probability space with filtration $\{\mathcal{F}(t) : t \geq 0\}$ is called a *stopping time* with respect to $\{\mathcal{F}(t) : t \geq 0\}$ if $\{\tau \leq t\} \in \mathcal{F}(t)$, for every $t \geq 0$.

Theorem 2.9. *Suppose that $\{W(t) : t \geq 0\}$ is a Brownian motion starting at $x \in \mathbb{R}$. For any $s > 0$, the process $\{W(t+s) - W(s) : t \geq 0\}$ is a Brownian motion started in the origin and is independent of the process $\{W(t) : 0 \geq t \geq s\}$.*

Proof. We omit the proof that the process above is a Brownian motion, which can be derived from the properties enumerated in Definition 2.2. The nontrivial claim that this new process is independent of $W(t)$ follows from the independent increments of Brownian motion, or intuitively that the change in the Brownian motion for $t > s$ is independent of $W(u)$, for $u \leq s$. □

Theorem 2.10 (Strong Markov Property)**.** *For every almost surely finite stopping time τ, meaning that $\mathbb{P}(\tau < \infty) = 1$, the process $\{W(\tau + t) - W(\tau) : t \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}^+(\tau) := \{A \in \mathcal{A} : A \cap \{\tau \leq t\} \in \mathcal{F}^+(t)$ for all $t \geq 0\}$.*

In essence, this statement builds on Theorem 2.9 and tells us that we can construct a Brownian motion with a *random* starting point from a standard Brownian motion, provided that the starting point is adapted to the Brownian motion, implied by the assumption that τ is a stopping time. Intuitively, this property tells us that even with stochastic time, Brownian motion is independent of its past history. In particular, it means that the Brownian motion after time τ relies only on $W(\tau)$.

Definition 2.11. A stochastic process $\{X(t) : t \geq 0\}$ defined on a filtered probability space, or a probability space along with a filtration, with filtration $\{\mathcal{F}(t) : t \geq 0\}$ is called *adapted* if $X(t)$ is $\mathcal{F}(t)-$measurable for any $t \geq 0$.

Definition 2.12. A real valued stochastic process $\{X(t) : t \geq 0\}$ is a *martingale* with respect to a filtration $\{\mathcal{F}(t) : t \geq 0\}$ if it is adapted to the filtration, $\mathbb{E}(|X(t)|) < \infty$ for all $t \geq 0$, and for any pair of times $0 \leq s \leq t$,

$$\mathbb{E}\left(X(t) \mid \mathcal{F}(s)\right) = X(s), \text{ almost surely.}$$

Morally, martingales describe *fair games*, in that if $X(s)$ represents the player's winnings in a particular game, the player is not expected to win or lose any more as the game progresses. Conversely, a game that is biased in favor of the house would have $\mathbb{E}(X(t) \mid \mathcal{F}(s)) < X(s)$, as the player is expected to lose more money as time goes on.

Before we show that Brownian motions are martingales, we must prove the following lemma.

Lemma 2.13. *Define $\mathcal{F}^+(s) = \bigcap_{t>s} \mathcal{F}^0(t)$, where $\mathcal{F}^0(t) = \sigma(W(s) : 0 \leq s \leq t)$ is the sigma algebra generated by the Brownian motion up to time t. For every $s \geq 0$ the process $\{W(t+s) + W(s) : t \geq 0\}$ is independent of the σ-algebra $\mathcal{F}^+(s)$*

Proof. Since Brownian motion has continuous paths, almost surely,

$$W(t+s) - W(s) = \lim_{n\to\infty} W(s_n + t) - W(s_n)$$

for a strictly decreasing sequence (s_n) converging to s. By the Markov Property, it follows that

$$\lim_{j\uparrow\infty}(W(t_1+s_j)-W(s_j),\ldots,W(t_m+s_j)-W(s_j))$$

is independent of $\mathcal{F}^+(s)$. Evaluating the limit,

$$(W(t_1+s)-W(s),\ldots,W(t_m+s)-W(s))$$

must be independent of the filtration, implying that the process

$$\{W(t+s)-W(s): t\geq 0\}$$

is also independent of $\mathcal{F}^+(s)$. □

Theorem 2.14. *Brownian motion is a martingale*

Proof. Take $\mathbb{E}(W(t)\mid\mathcal{F}^+(s))=\mathbb{E}(W(t)-W(s)\mid\mathcal{F}^+(s))+W(s)$. Since Brownian motion is independent of the filtration $\mathcal{F}^+(s)$,

$$\mathbb{E}(W(t)-W(s)\mid\mathcal{F}^+(s))+W(s)=\mathbb{E}(W(t)-W(s))+W(s).$$

Finally, since the increment is normally distributed with mean zero, $\mathbb{E}(W(t)-W(s))=0$. Thus, $\mathbb{E}(W(t)\mid\mathcal{F}^+(s))=W(s)$. □

So far, we primarily discussed one-dimensional Brownian motion for simplicity as these properties extend to higher dimensions. In this paper, however, we seek to prove the $n-$dimensional stochastic existence and uniqueness theorem, so we will provide the definition for $n-$dimensional Brownian motion below.

Definition 2.15. An $\mathbb{R}^n$-valued stochastic process $\mathbf{W}(\cdot)=(W_1(\cdot),\ldots,W_n(\cdot))$ is an *$n-$dimensional Brownian motion* if

(1) for each $k=1,\ldots,n$, $W_k(\cdot)$ is a one-dimensional Brownian motion and
(2) for $k=1,\ldots,n$, the $\sigma-$algebras generated by Brownian motions W_k denoted $\mathcal{W}_k=\sigma(W_k(t)\cdot t\geq 0)$ are independent.

3. Itô Calculus

Revisiting Equation (1.2), we interpret the differential equation in its integrated form:

$$X(t)=X_0+\int_0^t b(X(s),s)ds+\int_0^t \sigma(X(s),s)dW_s.$$

Before we can discuss this solution, we must first understand what it means to integrate with respect to a random process and the stochastic integral:

$$\int_0^t G(s)dW_s.$$

First, we must make sense of dW_t. We may consider $dW_t^2\approx dt$. One way to see this is that $W(t+h)-W(t)\sim\mathcal{N}(0,h)$. Then, $\mathbb{E}((W(t+h)-W(t))^2)=h$, implying $\mathbb{E}\left(\frac{(W(t+h)-W(t))^2}{h}\right)=1$. Sending h to zero, we see that

$$\mathbb{E}\left(\frac{dW_t^2}{dt}\right)=1.$$

To formalize this intuition, we will build the definition of the Itô integral like the construction of the Riemann-Stieltjes integral.

Definition 3.1. A *partition* P of an interval $[0, T]$ is a finite collection of points $t_i \in [0, T]$ such that

$$P := \{0 = t_0 < t_1 < \cdots < t_n = T\}.$$

Recall that when taking the Riemann integral of some function f, we approximate it with step functions $f_n(t) = f(s_j)$, $t_{j-1} \leq t \leq t_j$, where $t_{j-1} \leq s_j \leq t_j$, as follows:

$$\int_0^t f_n(s)ds = \sum_{i=1}^{n} f(s_j)(t_j - t_{j-1}).$$

We define the Riemann integral as the limit of the integrals of this sequence of step functions where the mesh size of the partition gets arbitrarily small, which is independent of the partitions and step functions that we choose:

$$\int_0^t f(s)ds = \lim_{n\to\infty} \int_0^t f_n(s)ds.$$

We construct the Itô integral in an identical way, but instead of approximating a stochastic process with step functions, we use the stochastic analog of *simple processes*, or a process that takes a finite number of values. Finally, we show that any function can be approximated by simple processes, allowing us to pass to limits.

Definition 3.2. A stochastic process $G(\cdot)$ is a *simple process* if there exist times $0 = t_0 < t_1 < \cdots < t_n = t$ and $\mathcal{F}(t_j)-$measurable random variables, G_k such that

$$G(t) = G_k, \qquad t_{k-1} \leq t < t_k.$$

Now, we define the stochastic integral as follows

Definition 3.3. Suppose $G(\cdot)$ is a simple process as defined above. Then,

$$\int_0^t G(s)dW_s := \sum_{k=1}^{n} Y_k(W(t_k) - W(t_{k-1})).$$

From this definition, we prove three important properties of the integration of simple processes, which easily extend to the general case.

Theorem 3.4. *Suppose $G(\cdot)$ and $H(\cdot)$ are simple processes. Then,*

(1) *(Linearity)*

$$\int_0^t aG(s) + bH(s)dW_s = a\int_0^t G(s)dW_s + b\int_0^t H(s)dW_s,$$

(2)

$$\mathbb{E}\left(\int_0^t G(s)dW_s\right) = 0, \ \textit{and}$$

(3) *(Itô Isometry)*

$$\mathbb{E}\left(\left(\int_0^t G(s)dW_s\right)^2\right) = \mathbb{E}\left(\int_0^t G^2(s)dt\right).$$

Proof. (1) The first statement can be proved similarly to the linearity of the Lebesgue integral, so we omit the proof.

(2) Suppose that G is a simple process where $G(t) = G_k$ for $t_{k-1} \le t < t_k$, Then,

$$\mathbb{E}\left(\int_0^t G(s)dW_s\right) = \mathbb{E}\left(\sum_{j=1}^n G_k(W(t_k) - W(t_{k-1}))\right).$$

We know that G_k is $\mathcal{F}(t_k)-$measurable and $\mathcal{F}(t_k)$ is independent of $\mathcal{W}^+(t_k)$. Since $W(t_k) - W(t_{k-1})$ is $\mathcal{W}^+(t_k)-$measurable, G_k is independent of $W(t_k) - W(t_{k-1})$. Therefore,

$$\begin{aligned}\mathbb{E}\left(\int_0^t G(s)dW_S\right) &= \mathbb{E}\left(\sum_{k=1}^n G_k(W(t_k) - W(t_{k-1})\right)\\ &= \sum_{k=1}^n \mathbb{E}(G_k)\mathbb{E}(W(t_k) - W(t_{k-1}))\\ &= \sum_{k=1}^n \mathbb{E}(G_k)\cdot 0 = 0.\end{aligned}$$

(3) Let G and G_k be as above. Then,

$$\mathbb{E}\left(\left(\int_0^t G(s)dW_s\right)^2\right) = \sum_{j=1}^n\sum_{k=1}^n \mathbb{E}(G_jG_k(W(t_j) - W(t_{j-1}))(W(t_k) - W(t_{k-1}))).$$

For all $j < k$, $W(t_j) - W(t_{j-1})$ is independent of $W(t_k) - W(t_{k-1})$ by the independent increments of Brownian motion, so $G_jG_k(W(t_j)-W(t_{j-1}))$ is independent of $W(t_k) - W(t_{k-1})$. Therefore, for $j < k$,

$$\begin{aligned}&\mathbb{E}(G_jG_k(W(t_j) - W(t_{j-1}))(W(t_k) - W(t_{k-1})))\\ =&\,\mathbb{E}(G_jG_k(W(t_j) - W(t_{j-1})))\mathbb{E}((W(t_k) - W(t_{k-1}))) = 0.\end{aligned}$$

This leaves the terms where $k = j$. Here we again use the independence of G_k from $W(t_k) - W(t_{k-1})$ to get

$$\begin{aligned}\mathbb{E}\left(\left(\int_0^t G(s)dW_s\right)^2\right) &= \sum_{k=1}^n \mathbb{E}(G_k^2(W(t_k) - W(t_{k-1}))^2)\\ &= \sum_{k=1}^n \mathbb{E}(G_k^2)\mathbb{E}((W(t_k) - W(t_{k-1}))^2)\\ &= \sum_{k=1}^n \mathbb{E}(G_k^2)(t_k - t_{k-1})\\ &= \mathbb{E}\left(\int_0^t G^2 dt\right).\end{aligned}$$

This final identity is especially important as it connects stochastic integration to deterministic integration, and we will leverage the Itô isometry in Section 5. □

Definition 3.5. $\mathbb{L}^2(0,T)$ is the space of all real-valued, progressively measurable stochastic processes $G(\cdot)$ such that

$$\mathbb{E}\left(\int_0^T G^2 ds\right) < \infty.$$

Any process $G \in \mathbb{L}^2(0,T)$ can be approximated by simple processes. Below, we will prove a weaker proposition, requiring the process to be continuous and uniformly bounded, but a sketch of the proof for the above statement can be found in [3] on page 67.

Lemma 3.6. *Let $G(\cdot)$ be a continuous process adapted to the filtration $\mathcal{F}$ such that, with probability one, $|G(t)| \leq C$, where $C < \infty$, for all $0 \leq t \leq T$. Then, there exists a sequence of bounded simple processes $G_n(\cdot)$ such that*

$$\lim_{n\to\infty} \int_0^t \mathbb{E}\left(|G_n(s) - G(s)|^2\right) ds \longrightarrow 0 = t.$$

Proof. Define

$$G_n(t) := n \int_{(j-1)/n}^{j/n} G(s) ds.$$

Then, by continuity, $G_n(t)$ converges almost surely to $G(t)$ uniformly. Now set

$$Y_n = \int_0^t |G_n(s) - G(s)|^2 ds.$$

Since all G_n are uniformly bounded in t, it follows from the dominated convergence theorem that

$$\lim_{n\to\infty} Y_n = \int_0^t \lim_{n\to\infty} |G_n(s) - G(s)|^2 ds = 0.$$

Since all Y_n are uniformly bounded as for any $n \in \mathbb{N}$,

$$Y_n \leq \int_0^t (|G_n(s)| + |G(s)|)^2 ds \leq 4TC^2,$$

we again apply the dominated convergence theorem to get that

$$\lim_{n\to\infty} \mathbb{E}(Y_n) = \mathbb{E}\left(\lim_{n\to\infty} Y_n\right) = 0.$$

□

As a result of the stronger version of this Lemma, holding for all continuous, adapted processes $G(\cdot) \in \mathbb{L}^2(0,T)$, we can now define the general stochastic integral.

Definition 3.7. Suppose $G(\cdot) \in \mathbb{L}^2(0,T)$,

$$\int_0^t G(s) dW_s = \lim_{n\to\infty} \int_0^t G_n(s) dW_s.$$

This definition is well defined as any two step processes approximating $G(\cdot)$ converge to zero, which can be seen by applying Itô isometry and then the Cauchy-Schwartz inequality.

Finally, we get to Itô's chain rule or Itô's lemma, which is the stochastic analog to the fundamental theorem of calculus and the chain rule. Critical in the proof of Itô's chain rule is Itô's product rule, below.

Theorem 3.8 (Itô's Product Rule)**.** *Suppose $X_1(\cdot)$ and $X_2(\cdot)$ are stochastic processes such that*

$$\begin{cases} dX_1 = F_1 dt + G_1 dW_t & (0 \leq t \leq T) \\ dX_2 = F_2 dt + G_2 dW_t & (0 \leq t \leq T), \end{cases}$$

where $F_i \in \mathbb{L}^1(0,T)$ and $G_i \in \mathbb{L}^2(0,T)$. Then,

$$d(X_1 X_2) = X_2 dX_1 + X_1 dX_2 + G_1 G_2 dt.$$

Remark 3.9. Analogous to Definition 3.5, $\mathbb{L}^1(0,T)$ is the space of all real-valued, progressively measurable stochastic processes F such that $\mathbb{E}(\int_0^t |F(s)| ds < \infty)$.

Notice that Itô's product rule is nearly identical to the deterministic product rule but with the addition of *Itô's correction term* at the end. The complete proof can be found in [3] on page 73.

Theorem 3.10 (Itô's Chain Rule)**.** *Suppose that a stochastic process $X(\cdot)$ has the stochastic differential*

$$dX = F dt + G dW_t,$$

where $F \in \mathbb{L}^1(0,T)$ and $G \in \mathbb{L}^2(0,T)$. Assume $u\colon \mathbb{R} \times [0,T] \longrightarrow \mathbb{R}$ is continuous and that $\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}$ exist and are continuous. Define

$$Y(t) := u(X(t), t).$$

Then,

$$\begin{aligned} dY &= \frac{\partial u}{\partial t} dt + \frac{\partial u}{\partial x} dX + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} G^2 dt \\ &= \left(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} F + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} G^2 \right) dt + \frac{\partial u}{\partial x} G dW_t. \end{aligned}$$

Proof. We begin by using induction along with Itô's product rule to prove that Itô's chain rule holds for polynomials. Then, we use the integrated form of the Stone-Weierstrass theorem, which states that we can approximate in the uniform norm any continuous function (and therefore any differentiable function) on a closed interval with polynomials. After establishing *uniform* convergence through Stone-Weierstrass, we pass to limits.

Let X be as above[3]. We claim that

$$d(X^m) = m X^{m-1} dX + \frac{1}{2} m(m-1) X^{m-2} G^2 dt.$$

The cases where $m = 0, 1$ trivially follow from the definition of dX. The case $m = 2$ follows from the Itô Product formula, where

$$d(X \cdot X) = 2X dX + G^2 dt.$$

Now, suppose our statement holds for $m - 1:$

$$\begin{aligned} d(X^{m-1}) &= (m-1) X^{m-2} dX + \frac{1}{2} m(m-1) X^{m-2} G^2 dt \\ &= \left((m-1) X^{m-2} F + \frac{1}{2} (m-1)(m-2) X^{m-3} G^2 \right) dt + (m-1) X^{m-2} G dW_t \end{aligned}$$

[3]To aid with the legibility of the proof, we drop the argument of $X(\cdot)$.

Then,

$$\begin{aligned}
d(X^m) &= d(X \cdot X^{m-1}) \\
&= XdX^{m-1} + X^{m-1}dX + (m-1)X^{m-2}G^2dt \\
&= X[((m-1)X^{m-2}F + \frac{1}{2}(m-1)(m-2)X^{m-3}G^2)dt \\
&+ (m-1)X^{m-2}GdW_t] + X^{m-1}dX + (m-1)X^{m-2}G^2dt \\
&= ((m-1)X^{m-1} + X^{m-1})dX \\
&+ \left(\frac{1}{2}(m-1)(m-2)X^{m-2}G^2 + (m-1)X^{m-2}G^2\right)dt \\
&= mX^{m-1}dX + \frac{1}{2}m(m-1)X^{m-2}G^2dt.
\end{aligned}$$

Since the differential operator is linear, we can apply Itô's formula to all polynomials u in X. Now, we want to extend that definition to include polynomials in both X and t. Suppose that $u(X,t) = f(X)g(t)$, where f and g are polynomials. Then,

$$\begin{aligned}
d(u(X,t)) &= d(f(X)g(t)) \\
&= f(X)dg(t) + gdf(X) \\
&= f(X)g'dt + g(f'(X)dX + \frac{1}{2}f''(X)G^2dt) \\
&= \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial x}dX + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}G^2dt.
\end{aligned}$$

Then, the statement above holds for any polynomial in X and t.

Now, for any given function u, satisfying the assumption of Itô's formula, there exists a sequence of polynomials (u_n) such that

$$u_n \longrightarrow u, \frac{\partial u_n}{\partial t} \longrightarrow \frac{\partial u}{\partial t}, \frac{\partial u_n}{\partial x} \longrightarrow \frac{\partial u}{\partial x}, \frac{\partial^2 u_n}{dx^2} \longrightarrow \frac{\partial^2 u}{\partial x^2}$$

uniformly on compact subsets of $\mathbb{R} \times [0,T]$.

After verifying the formula for polynomials, we know then that almost surely

$$u_n(X(r),r) - u_n(X(0),0) = \int_0^r \frac{\partial u_n}{\partial t} + \frac{\partial u_n}{\partial x} + \frac{1}{2}\frac{\partial^2 u_n}{\partial x^2}G^2dt + \int_0^r \frac{\partial u_n}{\partial x}GdW_t$$

for all $0 \leq r \leq T$. Passing to limits as (u_n) converge uniformly, we prove Itô's chain rule in one dimension. □

4. The Deterministic Existence and Uniqueness Theorem

Now, we return to our consideration of deterministic differential equations, as the proof of the existence of a unique solution extends nicely to its stochastic counterpart. Integrating the differential form of Equation (1.1), we know that our solution must be some path $x(\cdot)$ that satisfies

$$x(t) = x_0 + \int_0^t b(x(s))ds.$$

The classical way of proving such path exists is by using *Picard iteration*, where we inductively define a sequence of functions and show that they must converge to our solution. Then, we implement *Gronwall's inequality* to show that the given solution is unique. An alternate proof that can also extend to our stochastic equation uses the *Banach fixed point theorem* and can be found in [5].

Lemma 4.1 (Gronwall)**.** *If* $\phi(t) \leq a + b\int_0^t \phi(s)ds$*, where* $a \geq 0$*,* $b > 0$*, then* $\phi(t) \leq ae^{bt}$*. In particular, if* $a = 0$ *and* $\phi \geq 0$*,* $\phi \equiv 0$*.*

Proof. Let $\Phi(t) = \int_0^t \phi(s)ds$. Then, by our assumption above, $\Phi'(t) \leq a + b\Phi(t)$. Multiplying by the integrating factor e^{-bt},

$$(\Phi e^{-bt})' = \Phi'(t)e^{-bt} - b\Phi(t)e^{-bt} \leq ae^{-bt}.$$

Noting that $\Phi(0) = \int_0^0 \phi(s)ds = 0$, we integrate to get

$$\begin{aligned}\Phi(t)e^{-bt} - \Phi(0) &\leq \frac{-a}{b}e^{-bt} + \frac{a}{b}\\ \Phi(t)e^{-bt} &\leq \frac{-a}{b}(e^{-bt} - 1)\\ \Phi(t) &\leq \frac{-a}{b}(1 - e^{bt}).\end{aligned}$$

Finally,

$$\phi(t) \leq a + b\Phi(t) \leq a - a + ae^{bt} = ae^{bt}.$$

□

Theorem 4.2. *Suppose* $b : \mathbb{R}^d \times [0,T] \longrightarrow \mathbb{R}^d$ *is globally Lipschitz continuous, that is* $|b(x,t) - b(y,t)| \leq M|x - y|$ *for all* $x, y \in \mathbb{R}^d$*,* $M \in \mathbb{R}$*. Then, the deterministic ordinary differential equation*

$$\begin{cases}\dot{x}(t) = b(x(t),t) & t > 0\\ x(0) = x_0,\end{cases}$$

has a unique solution.

Proof. (1) *Existence.*

We first define continuous $x_i \in X := C\left([0,\delta], \mathbb{R}^d\right)$, the space of continuous (vector-valued) paths on $[0,\delta]$, equipped with the uniform norm defined as

$$\|x - y\|_X = \sup_{0 \leq t \leq \delta} |x(t) - y(t)|.$$

We set $x_0(t) = x_0$, and inductively define

$$x_{n+1}(t) = x_0 + \int_0^t b(x_n(s),s)ds, \qquad t \in [0,\delta].$$

We now aim to show that this sequence must be Cauchy and therefore, will converge in the complete space, X. Take $|x_{n+1}(t) - x_n(t)|$. Using the definition of x_i, the Cauchy-Schwartz inequality, and the Lipschitz continuity of b, we get

$$\begin{aligned}|x_{n+1}(t) - x_n(t)| &= \int_0^t b(x_n(s),s)ds - \int_0^t b(x_{n-1}(s),s)ds\\ &\leq \int_0^t |b(x_n(s),s) - b(x_{n-1}(s),s)|ds\\ &\leq \int_0^t M\|x_n - x_{n-1}\|ds\\ &\leq \delta M\|x_n - x_{n-1}\|.\end{aligned}$$

If we choose δ so that $\delta M < \frac{1}{2}$,

$$\|x_{n+1} - x_n\| \leq \frac{1}{2}\|x_n - x_{n-1}\|.$$

Iterating, we get $\|x_{n+1} - x_n\| \leq 2^{-n+1}\|x_1 - x_0\|$. We now use inequality to prove that the sequence is Cauchy. Let $C = \|x_1 - x_0\|$ and suppose $n \geq m$. Then,

$$\begin{aligned}\|x_n - x_m\| &\leq \sum_{k=m}^{n-1} \|x_{k+1} - x_k\| \\ &\leq C \sum_{k=m}^{n-1} 2^{-k+1} \\ &\leq C \sum_{k=m}^{\infty} 2^{-k+1}\end{aligned}$$

Since the series converges, this sequence must be Cauchy, so (x_n) converges uniformly to some path $x(\cdot)$. Thus, by the uniform convergence theorem, we pass to limits through the integral to find

$$x(t) = x_0 + \int_0^t b(x(s), s)ds.$$

Since $b(\cdot)$ is smooth and $x(\cdot)$ is continuous, the right hand side is differentiable by the fundamental theorem of calculus. Therefore, $x(\cdot)$ is differentiable. Differentiating, we see that x is a classical solution:

$$\dot{x}(t) = b(x(s), s).$$

(2) *Uniqueness.*

Suppose x and y are solutions with the same initial condition. Then, again using the Lipschitz property of b,

$$|x(t) - y(t)| = \int_0^t b(x(s), s) - b(y(s), s)ds \leq M \int_0^t |x(s) - y(s)|ds.$$

If $\phi = |x(t) - y(t)|$, we can see that the inequality above satisfies the assumption of Gronwall's inequality with $a = 0$ and $b = M$. Thus, $|x(t) - y(t)| \equiv 0$, implying that our solution must be unique. □

5. The Stochastic Existence and Uniqueness Theorem

Before we introduce the stochastic Existence and Uniqueness Theorem, we state a lemma vital to the proof.

Lemma 5.1 (Borel-Cantelli)**.** *Suppose (A_n) is a sequence of events. If*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$$

then,

$$\mathbb{P}(A_n \textit{ infinitely often}) = 0.$$

Proof. This is a critical lemma in probability theory, the proof of which can be found in [3] on page 20. □

Theorem 5.2 (Existence and Uniqueness)**.** *Suppose that* $b : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ *and* $\sigma : \mathbb{R}^n \times [0,T] \to M^{m\times n}$ *are globally uniformly Lipschitz continuous, namely that*

$$|b(x,t) - b(y,t)| \le L|x-y|$$
$$|\sigma(x,t) - \sigma(y,t)| \le L|x-y|$$

for all $x, y \in \mathbb{R}^n$ *and* $0 \le t \le T$, *where* L *is some constant.*

Next, let X_0 *be any* $\mathbb{R}^n-$*valued random variable such that*

$$\mathbb{E}(|X_0|^2) < \infty$$

and

$$X_0 \textit{ is independent of } \sigma(0),$$

where $W(\cdot)$ *is a given* $m-$*dimensional Brownian motion.*

Then, almost surely, there exists a unique solution $X \in \mathbb{L}^2_n(0,T)$ *of the stochastic differential equation in the form of Equation* (1.2).

Proof. (1) *Existence.* We begin by inductively defining a sequence of random variables, and then show that they uniformly converge in probability to a solution, as in the deterministic case.

Let $X_0(t) = X_0$, and define

$$X_{n+1}(t) = X_0 + \int_0^t b(X_n(s), s)ds + \int_0^t \sigma(X_n(s), s)dW_s.$$

Like the proof of Theorem 4.2, we wish to prove that the sequence (X_n) is almost surely Cauchy. First, we note from the definition above that for any $n \in \mathbb{N}$,

$$X_n(t) = X_0 + \sum_{j=0}^{n-1}(X_{j+1}(t) - X_j(t)),$$

and thus, for $n \ge m$

$$X_n(t) - X_m(t) = \sum_{j=n-m}^{n}(X_{j+1}(t) - X_j(t)).$$

We aim to show that the difference between two consecutive terms for a sufficiently large enough n is arbitrarily small or that

$$\mathbb{P}\left(\max_{0\le t\le T}|X_{n+1}(t) - X_n(t)| > \frac{1}{2^n} \text{ infinitely often}\right) = 0.$$

We will prove the statement above by showing that the assumption of the Borel-Cantelli Lemma holds. To do so, we will use Doob's martingale inequality, proved in the appendix.

We first claim that

$$\mathbb{E}(|X_{n+1}(t) - X_n(t)|^2) \le \frac{(Mt)^{n+1}}{(n+1)!},$$

which we will prove by induction. First, we know that since b is uniformly Lipschitz, $|b(x,t)| \le C(1+|x|)$, where $C := \max\{b(0,t), L\}$. This inequality similarly applies to σ. The short proof is as follows:

$$|b(x,t)| \le |b(x,t) - b(0,t)| + |b(0,t)| \le L|x| + |b(0,t)| \le C(|x|+1).$$

With this inequality, we consider our base case

$$\mathbb{E}(|X_1(t) - X_0(t)|^2) = \mathbb{E}\left(\left|\int_0^t b(X_0(s), s)ds + \int_0^t \sigma(X_0(s), s)dW_s\right|^2\right).$$

Using the inequality $(a+b)^2 \leq 2a^2 + b^2$ along with Itô Isometry,

$$\begin{aligned}
\mathbb{E}(|X_1(t) - X_0(t)|^2) &= \mathbb{E}\left(\left|\int_0^t b(X_0(s), s)ds + \int_0^t \sigma(X_0(s), s)dW_s\right|^2\right) \\
&\leq 2\mathbb{E}\left(\left|\int_0^t b(X_0(s), s)ds\right|^2\right) + 2\mathbb{E}\left(\left|\int_0^t C(|X_0(s)| + 1)dW_s\right|^2\right) \\
&\leq 2\mathbb{E}\left(\left|\int_0^t C(|X_0(s)| + 1)ds\right|^2\right) + 2\mathbb{E}\left(\int_0^t C^2(|X_0(s)| + 1)^2 ds\right) \\
&\leq Mt,
\end{aligned}$$

where M is some large constant. In this last step, we rely on the assumption that $\mathbb{E}(|X_0^2|)$ is finite as well as the fact that when integrating a random variable with a deterministic integral, we treat the random variable as a constant.

Now, assume $\mathbb{E}(|X_n(t) - X_{n-1}(t)|^2) \leq \frac{M^n t^n}{n!}$. Then, applying the same identity as before along with the Cauchy-Schwartz inequality

$$\begin{aligned}
\mathbb{E}(|X_{n+1}(t) - X_n(t)|^2) &= \mathbb{E}\left(\left|\int_0^t b(X_n(s), s) - b(X_{n-1}(s), s)ds + \int_0^t \sigma(X_n(s), s) - \sigma(X_{n-1}(s), s)dW\right|^2\right) \\
&\leq 2\mathbb{E}\left(\left|\int_0^t b(X_n(s), s) - b(X_{n-1}(s), s)ds\right|^2\right) \\
&+ 2\mathbb{E}\left(\left|\int_0^t \sigma(X_n(s), s) - \sigma(X_{n-1}(s), s)dW\right|^2\right) \\
&\leq 2T\mathbb{E}\left(\int_0^t L^2|X_n(s) - X_{n-1}(s)|^2 ds\right) + 2\mathbb{E}\left(\int_0^t |\sigma(X_n(s), s) - \sigma(X_{n-1}(s), s)|^2 ds\right) \\
&\leq 2TL^2\mathbb{E}\left(\int_0^t |X_n(s) - X_{n-1}(s)|^2 ds\right) + 2L^2\mathbb{E}\left(\int_0^t |X_n(s) - X_{n-1}(s)|^2 ds\right).
\end{aligned}$$

By Fubini's theorem, allowing us to interchange integrals, and finally applying our induction hypothesis

$$\begin{aligned}
\mathbb{E}(|X_{n+1}(t) - X_n(t)|^2) &\leq 2TL^2\mathbb{E}\left(\int_0^t |X_n(s) - X_{n-1}(s)|^2 ds\right) + 2L^2\mathbb{E}\left(\int_0^t |X_n(s) - X_{n-1}(s)|^2 ds\right) \\
&= 2TL^2\left(\int_0^t \mathbb{E}(|X_n(s) - X_{n-1}(s)|^2)ds\right) + 2L^2\left(\int_0^t \mathbb{E}(|X_n(s) - X_{n-1}(s)|^2)ds\right) \\
&\leq 2L^2(T+1)\left(\int_0^t \mathbb{E}(|X_n(s) - X_{n-1}(s)|^2)ds\right) \\
&\leq 2L^2(T+1)\left(\int_0^t \frac{M^n s^n}{n!}ds\right) \\
&\leq \frac{M^{n+1}t^{n+1}}{(n+1)!},
\end{aligned}$$

provided we choose $M \geq 2L^2(1+T)$. With that, we have proven our claim.

Now, consider that

$$\begin{aligned}\max_{0\leq t\leq T}|X_{n+1}(t)-X_n(t)|^2 &\leq 2\max_{0\leq t\leq T}\left|\int_0^t b(X_n(s),s)-b(X_{n-1}(s),s)ds\right|^2 \\ &+2\max_{0\leq t\leq T}\left|\int_0^t \sigma(X_n(s),s)+\sigma(X_{n-1}(s),s)dW_s\right|^2 \\ &\leq 2TL^2\int_0^t |X_n(s)-X_{n-1}(s)|^2 ds \\ &+2\max_{0\leq t\leq T}\left|\int_0^t \sigma(X_n(s),s)+\sigma(X_{n-1}(s),s)dW_s\right|^2.\end{aligned}$$

We then take the expectation of both sides of the inequality and apply our result above along with the the martingale inequality to get

$$\begin{aligned}\mathbb{E}\left(\max_{0\leq t\leq T}|X_{n+1}(t)-X_n(t)|^2\right) &\leq 2TL^2\int_0^t \mathbb{E}(|X_n(s)-X_{n-1}(s)|^2)ds \\ &+8\mathbb{E}\left(\left|\int_0^t \sigma(X_n(s),s)+\sigma(X_{n-1}(s),s)dW_s\right|^2\right) \\ &\leq 2TL^2\int_0^t \mathbb{E}(|X_n(s)-X_{n-1}(s)|^2)ds \\ &+8L^2\int_0^t \mathbb{E}(|X_n(s)-X_{n-1}(s)|^2)ds \\ &\leq C\frac{M^nT^n}{n!}.\end{aligned}$$

Applying Chebyshev's inequality

$$\begin{aligned}\mathbb{P}\left(\max_{0\leq t\leq T}|X_{n+1}(t)-X_n(t)|>\frac{1}{2^n}\right) &\leq 2^{2n}\mathbb{E}\left(\max_{0\leq t\leq T}|X_{n+1}(t)-X_n(t)|^2\right) \\ &\leq 2^{2n}C\frac{M^nT^n}{n!}\end{aligned}$$

Since

$$\sum_{n=1}^{\infty} 2^{2n}C\frac{M^nT^n}{n!}<\infty,$$

$$\sum_{n=1}^{\infty}\mathbb{P}\left(\max_{0\leq t\leq T}|X_{n+1}(t)-X_n(t)|>\frac{1}{2^n}\right)<\infty.$$

Therefore, applying Borel-Cantelli lemma, we know that

$$\mathbb{P}\left(\max_{0\leq t\leq T}|X_{n+1}(t)-X_n(t)|>\frac{1}{2^n}\quad\text{infinitely often}\right)=0,$$

and thus, the sequence must converge uniformly on $[0,T]$ to some process $X(t)$.

Passing to limits in our definition of X_{n+1}, we prove there exists

$$X(t)=X_0+\int_0^t b(X(s),s)ds+\int_0^t \sigma(X(s),s)dW_s,$$

and in differential form

$$\begin{cases} dX = b(X(t),t)dt + \sigma(X(t),t)dW_t \\ X(0) = X_0. \end{cases}$$

2. *Uniqueness.* Similar to the deterministic case, this proof will rely on Gronwall's inequality, and setting up the assumption for the inequality will be very similar to the techniques used for the proof of the existence.

Suppose that X and Y are solutions to the differential equation defined above. Then, we estimate

$$\begin{aligned} \mathbb{E}(|X(t)-Y(t)|^2) &\leq 2\mathbb{E}\left(\left|\int_0^t b(X(s),s)-b(Y(s),s)ds\right|^2\right) + 2\mathbb{E}\left(\left|\int_0^t \sigma(X(s),s)-\sigma(Y(s),s)dW_s\right|^2\right) \\ &\leq 2T\mathbb{E}(\int_0^t |b(X(s),s)-b(Y(s),s)|^2 ds) + 2\mathbb{E}(\int_0^t |\sigma(X(s),s)-\sigma(Y(s),s)|^2 ds) \\ &\leq 2L^2T\int_0^t \mathbb{E}(|X(t)-Y(t)|^2)ds + 2L^2\int_0^t \mathbb{E}(|X(t)-Y(t)|^2)ds \\ &\leq C\int_0^t \mathbb{E}(|X(t)-Y(t)|^2)ds, \end{aligned}$$

where $C \geq 2L^2(T+1)$.

This now satisfies the assumption of Gronwall's inequality with $a = 0$ and $b = C$. Therefore, $\mathbb{E}(|X(t)-Y(t)|^2) \equiv 0$, implying that

$$\mathbb{P}\left(\max_{0\leq t\leq T}|X(t)-Y(t)| > 0\right) = 0,$$

for all $X(t), Y(t)$, almost surely. Thus, our solution must be unique. □

6. Applications

Stochastic differential equations are useful in mathematical finance, and one way we can apply this is deriving the *Black-Scholes-Merton* PDE, which models the price of European call options.

A *derivative* is a financial instrument whose payoff depends on the value of an underlying security. A *call option* is a contract which allows the buyer the right to buy a share of the stock S at a *strike price* p at an *expiration time.* A *European call option* only allows the buyer to exercise the option at the expiration time.

Let $S(t)$ denote the price of a given security at time t, and suppose that S evolves according to a geometric Brownian motion, as discussed above:

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t \\ S(0) = s_0. \end{cases}$$

We position ourselves as a financial firm that wishes to sell a European call option of S at strike price p and expiration time T. Assume that the interest rate r is constant. We now ask what is appropriate price we should sell the option for.

While constructing an appropriate model, we want to create no opportunities for *arbitrage* or risk-free profits for others in the market. To do this, we will consider a hedging strategy where we *replicate* the option.

First, take $C(t) := u(S(t), t)$ to represent the price of the option at time t. Since it relies on a random variable, $C(t)$ is random, so we use Itô's formula to get the differential of C

$$dC(t) = \left(\frac{\partial C}{\partial t} + \mu S(t)\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 C}{\partial S^2}\right) dt + \sigma S(t)\frac{\partial C}{\partial S} dW_t.$$

We will replicate C by a portfolio, P with shares of our stock S and shares of our bond B, which we assume is a risk-free investment. Note that this implies that

$$\begin{cases} dB(t) = rB(t)dt \\ B(0) = 1, \end{cases}$$

and therefore, $B(t) = e^r t$. We represent the units of stock S by the random process $\phi(\cdot)$ and the units of stock B by $\psi(\cdot)$. So,

$$P(t) = \phi(t)S(t) + \psi(t)B(t) \ \ (0 \le t \le T).$$

The replicating assumption then implies that this is self financing or that

$$\begin{aligned} dP(t) &= \phi(t)dS_t + \psi(t)dB_t \\ &= (r\psi(t)B(t) + \phi(t)\mu S(t))\, dt + \phi(t)\sigma S(t)dW_t. \end{aligned}$$

We now wish to set dP to dC, so we see that

$$\begin{aligned} \phi_t &= \frac{\partial C}{\partial S} \\ r\psi_t B_t &= \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}. \end{aligned}$$

If we set $C_0 = P_0$, knowing that the dynamics of C and P are the same means that $C(t) = P(t)$ for all t. Thus,

$$rS(t)\frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 C}{\partial S^2} - rC(t) = 0.$$

The PDE above is the Black-Scholes-Merton PDE, and the solution is

$$C(S(t), t) = S(t)\Phi(d_1) - e^{-r(T-t)} K\Phi(d_2)$$

$$\text{where } d_1 = \frac{\log\left(\frac{S(t)}{K}\right) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t},$$

and Φ is the cumulative distribution function of the standard normal distribution.

From the Black-Scholes PDE, we can derive the Black-Scholes formula, which, with some adjustments, is widely used in the options market. The application above is just one of the many ways we can use stochastic differential equations to model the dynamics of financial markets.

7. Appendix: Martingale Inequality

We begin this section with three helpful lemmas.

Lemma 7.1 (Doob's Stopping Theorem or Optional Stopping Theorem)**.** *Let $\{\mathcal{F}(t) : t \geq 0\}$ be a filtration defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\{M(t) : t \geq 0\}$ be a martingale adapted to the above filtration with right continuous and locally bounded paths. Then, for any almost surely bounded stopping time τ of the filtration such that $\mathbb{E}(|M(t)|) < +\infty$,*

$$\mathbb{E}(M(\tau)) = \mathbb{E}(M(0)).$$

Lemma 7.2 (Jensen's Inequality)**.** *Let X be a random variable and g a convex function, then*

$$g(\mathbb{E}(X)) \leq E(g(X)).$$

Lemma 7.3 (Fubini's Theorem)**.** *Suppose a function f is integrable over $X \times Y$. Then,*

$$\int_{X\times Y} f(x,y)d\mu = \int_Y \int_X f(x,y)dxdy = \int_X \int_Y f(x,y)dydx.$$

Theorem 7.4 (Doob's Martingale Inequality)**.** *If $X(\cdot)$ is a martingale and $1 < p < \infty$, then*

$$\mathbb{E}\left(\max_{0\leq s\leq t} |X(s)|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(|X(t)|^p).$$

Proof. Let $p \geq 1$ and $T > 0$. If $\mathbb{E}(|M_T|^p) < +\infty$, by Jensen's inequality with $g(x) = x^p$, $\{|M_t|^p : 0 \leq t \leq T\}$ is a submartingale[4]. Let $\lambda > 0$ and $\tau = \inf\{s \geq 0 \mid |M_s| \geq \lambda\} \wedge T$, where $\wedge$ denotes the minimum. Here, we can see that τ is a stopping time, bounded by T. Then, applying the optional stopping theorem,

$$\mathbb{E}(|M(\tau)|^p) \leq \mathbb{E}(|M(T)|^p).$$

Denote $X := \sup_{0\leq t\leq T}|M(t)|$. By the definition of τ,

$$\mathbb{1}_{X\geq\lambda}\lambda^p + \mathbb{1}_{X<\lambda}|M(T)|^p \leq |M(\tau)|^p.$$

Then,

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|M(t)| \geq \lambda\right) \leq \frac{1}{\lambda^p}\mathbb{E}(|M(T)|^p \mathbb{1}_{X>\lambda}) \leq \frac{1}{\lambda^p}\mathbb{E}(|M(T)|^p).$$

Multiplying the inequality above by λ and integrating, we deduce that

$$\int_0^{+\infty} \lambda^{p-1}\mathbb{P}\left(\sup_{0\leq t\leq T}|M(t)| \leq \lambda\right) d\lambda \leq \int_0^{+\infty} \lambda^{p-2}\mathbb{E}(|M(T)|\mathbb{1}_{X\geq\lambda}d\lambda.$$

Applying Fubini's theorem,

$$\int_0^{+\infty} \lambda^{p-1}\mathbb{P}\left(\sup_{0\leq t\leq T}|M(t)| \leq \lambda\right) d\lambda = \int_\Omega \int_0^X \lambda^{p-1} d\lambda d\mathbb{P}(\omega) = \frac{1}{p}\mathbb{E}\left(\left(\sup_{0\leq t\leq T}|M(t)|\right)^p\right).$$

Similarly,

$$\int_0^{+\infty} \lambda^{p-2}\mathbb{E}(|M(T)|\mathbb{1}_{X\geq\lambda})d\lambda = \frac{1}{p-1}\mathbb{E}\left(\left(\sup_{0\leq t\leq T}|M(t)|\right)^{p-1}|M(T)|\right).$$

---

[4]If a process X is a submartingale, $\mathbb{E}(X(t) \mid \mathcal{F}(s)) \geq X(s)$.

Thus,

$$\mathbb{E}\left(\left(\sup_{0\le t\le T}|M(t)|\right)^p\right) \le \frac{p}{p-1}\mathbb{E}\left(\left(\sup_{0\le t\le T}|M(t)|\right)^{p-1}|M(T)|\right).$$

Using Hölder's inequality, we see that

$$E\left(\left(\sup_{0\le t\le T}|M(t)|\right)^{p-1}|M(T)|\right) \le \mathbb{E}(|M(T)|^p)^{\frac{1}{p}}\mathbb{E}\left(\left(\sup_{0\le t\le T}|M(t)|\right)^p\right)^{\frac{p-1}{p}}.$$

Then, if $\mathbb{E}((\sup_{0\le t\le T}|M(t)|)^p) < +\infty$, applying the optional stopping theorem once again,

$$\mathbb{E}\left(\left(\sup_{0\le t\le T}|M(t)|\right)^p\right) \le \left(\frac{p}{p-1}\right)^p \mathbb{E}(|M(T)|^p).$$

If $\mathbb{E}\left(\left(\sup_{0\le t\le T}|M(t)|\right)^p\right) = +\infty$, we consider for $N \in \mathbb{N}$ the stopping time $\tau_N = \inf\{t \ge 0 \mid |M(t)| \ge N\} \wedge T$. If we consider the above inequality instead with respect to the martingale $\{M(t\wedge\tau_N)\}$, we obtain

$$\mathbb{E}\left(\left(\sup_{0\le t\le T}|M(t\wedge\tau_N)|\right)^p\right) \le \left(\frac{p}{p-1}\right)^p \mathbb{E}(|M(T)|^p).$$

By the monotone convergence theorem, we get our desired result on the left hand side. □

Acknowledgments

It is my great pleasure to thank my mentor, David Bowman, without whose excellent teaching, guidance, and thoughtfulness, none of this would have been possible. I would especially like to thank him for pushing me to learn a great deal of advanced mathematics, and his patience with my inexperience. Additionally, I would like to thank Alexa Pomerantz for helping me polish this paper and revising my work. I would also like to thank Daniil Rudenko and Laszlo Babai for delivering the apprentice lectures and spending countless hours in problem sessions. Finally, I would like to extend my gratitude to Peter May for organizing the REU and giving me the opportunity to further my mathematical studies.

References

[1] Peter Morters and Yuval Peres. Brownian Motion. Cambridge University Press. 2010.
[2] Gregory F. Lawler. Stochastic Calculus: An Introduction with Applications. http://www.math.uchicago.edu/ lawler/finbook.pdf
[3] Lawrence C. Evans. An Introduction to Stochastic Differential Equations Version 1.2. https://www.cmor-faculty.rice.edu/ cox/stoch/SDE.course.pdf
[4] Martin Haugh. The Black-Scholes Model. http://www.columbia.edu/ mh2078/FoundationsFE/BlackScholes.pdf
[5] Fabrice Baudoin. Lecture 11. Doob's martingale maximumal inequalities. https://fabricebaudoin.wordpress.com/2012/04/10/lecture-11-doobs-martingale-maximal-inequalities/