Numerically Solving The Dirichlet Problem Using Stochastic Calculus

Leonardo Ferreira Guilhoto

Summer 2019

Abstract

In this expository paper we introduce stochastic calculus and some of the numerical methods used in this field. Namely, we introduce: the Itô integral; Itô's formula; stochastic differential equations (SDEs); Itô diffusions; stochastic Taylor expansions; and discrete-time approximations of SDEs (Euler-Maruyama and Milstein methods). We illustrate this material by using it to tackle the generalized Dirichlet problem, first relating it to Itô diffusions and then studying a specific numeric implementation by applying it to two example problems.

Contents

1 Introduction

This paper presents a survey of the basics of Stochastic Calculus, with a focus on numerical solutions to Stochastic Differential Equations (SDEs) and applications to the generalized Dirichlet problem. The target reader for this paper has a solid background on the basics of real analysis and some familiarity with probability. Knowledge of stochastic processes (especially Brownian Motions) is particularly useful, but a brief introduction of the material used is given in Section 1.1. It should be noted that despite there being two distinct definitions for stochastic calculus (one due to the Japanese mathematician Kiyoshi Itô and the other due to the Russian physicist Ruslan Stratonovich), this paper deals solely with Itô calculus, meaning that all stochastic integrals should be interpreted as presented in Section 3.1. The material presented in sections 1 through 4 is mostly expository while the particular implementation in Section 5 using the original code from Appendix A is my own[1].

1.1 Preliminaries

Before we delve into the specifics of Stochastic Calculus, we review some basic concepts from Probability Theory and present what is perhaps the most important stochastic process in mathematics: Brownian motion. For a more thorough explanation of the topics presented in this section, the reader may refer to any standard textbook on the matter such as [LL10]. If the reader is familiar with these topics, they may find it useful to skip to Section 2.

1.1.1 Probability and Measure Theory

Definition 1.1. Let Ω be a set. A σ-algebra on Ω is a family $\mathcal{F}$ of subsets of Ω satisfying:

- $\emptyset \in \mathcal{F}$.
- If $F \in \mathcal{F}$, then its complement F^C is also in F. That is, $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$.
- If $F_1, F_2, \ldots$ is a sequence of subsets of Ω included in $\mathcal{F}$, then their union is also in $\mathcal{F}$. That is, $F_1, F_2, \cdots \in \mathcal{F} \Rightarrow \cup_{i=1}^{\infty} F_i \in \mathcal{F}$.

Furthermore, we call a pair $(\Omega, \mathcal{F})$ where $\mathcal{F}$ is a σ-algebra on Ω a measurable space.

Note that the first and second requirements in this definition imply that if $\mathcal{F}$ is a σ-algebra then $\Omega \in \mathcal{F}$.

Definition 1.2. Let $(\Omega, \mathcal{F})$ be a measurable space. A probability measure on $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$ satisfying:

- $P(\emptyset) = 0$ and $P(\Omega) = 1$.
- If $F_1, F_2, \cdots \in \mathcal{F}$ and $F_i \cap F_j = \emptyset$ whenever $i \neq j$, then $P\left(\cup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} P(F_i)$.

Furthermore, we call the triple $(\Omega, \mathcal{F}, P)$ where P is a probability measure on $(\Omega, \mathcal{F})$ a probability space.

In order to gain some intuition we observe what is perhaps the simplest non-trivial probability space: the flipping of a fair coin.

Example 1.3. Consider the set $\Omega = \{H, T\}$ composed of the possible results of the flip of a coin (H representing heads and T representing tails). Additionally, let $\mathcal{F} = \left\{\emptyset, \{H\}, \{T\}, \Omega\right\}$, $P(\emptyset) = 0$, $P(\Omega) = 1$ and $P(\{H\}) = P(\{T\}) = \frac{1}{2}$. Then $(\Omega, \mathcal{F}, P)$ is the probability space that depicts the flipping of a fair coin. One can check that this satisfies all the requirements stated in definitions 1.1 and 1.2. One can then make the following interpretations:

- Ω represents the **set of all possible outcomes**, which in this case is the outcome of the flip of a coin: heads or tails.
- $\mathcal{F}$ represents the **set of possible events**, where an event is thought of as a subset of all possibilities. That is, an event is a collection of outcomes. For example, in the case above $\{H\}$ represents the event of the coin coming up heads, $\{T\}$ the event of the coin coming up tails, $\Omega = \{H, T\}$ the event of the coin coming up heads **or** tails and $\emptyset$ the event of the coin coming up **neither** heads **nor** tails.

[1] I do not claim to be the first person to think of using the Monte Carlo method to solve this type of problem, only that this particular implementation of it was designed and implemented by myself

- P represents the **probability of a given event happening**. Accordingly, we have that $P(\emptyset) = 0$, $P(\Omega) = 1$ and $P(\{H\}) = P(\{T\}) = \frac{1}{2}$.

Definition 1.4. Let $(\Omega, \mathcal{F}, P)$ be a probability space and (S, Σ) be a measurable space. We say a function $X : \Omega \to S$ is $\mathcal{F}$-measurable if for any subset $U \in \Sigma$ the pre-image of U is in $\mathcal{F}$, that is, if $X^{-1}(U) \in \mathcal{F}$ for any U in Σ. If X is a $\mathcal{F}$-measurable function we call it a random variable and say that its distribution is the function

$$\mu_X(U) := P\left(X^{-1}(U)\right),$$

which takes elements of Σ to $[0, 1]$. Furthermore, we call S the state space.

In this paper we take S to be $\mathbb{R}^n$ and Σ to be the the Borel σ-algebra $\mathcal{B}$, generated by all open subsets of $\mathbb{R}^n$. One can show that with these definitions the pair $(\mathbb{R}^n, \mathcal{B})$ is indeed a measure space.

Definition 1.5. Let $(\Omega, \mathcal{F}, P)$ be a probability space and T be totally ordered a set. A stochastic process is a parametrized collection of random variables $X(t) : \Omega \to \mathbb{R}^n$ defined for all $t \in T$. Additionally, for a fixed $\omega \in \Omega$ we say that the function $t \mapsto X(t)[\omega]$ is the path of $X(t)$ associated to ω.

In most cases, especially within this paper, T is taken to be the set of non-negative real numbers $[0, \infty)$ and we interpret t as time. From now on in the paper we assume that $T = [0, \infty)$.

Definition 1.6. Let $(\Omega, \mathcal{F})$ be a measurable space. A filtration is a family of σ-algebras $\mathcal{M} = \{\mathcal{M}_t\}_{t\geq 0}$ with $\mathcal{M}_t \subset \mathcal{F}$ for all t such that $0 \leq s \leq t$ implies $\mathcal{M}_s \subset \mathcal{M}_t$.

A filtration can be thought of as the evolution of a given stochastic process and $\mathcal{M}_t$ as the information contained in the history of that process up to and including time t.

Definition 1.7. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration of $\mathcal{F}$, and $X(t) : \Omega \to \mathbb{R}^n$ be a stochastic process. Then X is said to be an adapted process in regards to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if for all $s \geq 0$ $X(s)$ is $\mathcal{F}_s$-measurable.

That is, saying that a process is adapted means that it cannot see into the future. In other words, that all the information contained in $\mathcal{F}_t$ is enough to compute $X(t)$.

Definition 1.8. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{M}$ a filtration. A stochastic process $X(t) : \Omega \to \mathbb{R}^n$ which is adapted to $\mathcal{F}$ has the Markov property (also known as memoryless property) if for any s, t with $0 < s < t$ and any open set $U \subset \mathbb{R}^n$ we have that

$$P\left(X(t) \in U | \mathcal{F}_s\right) = P\left(X(t) \in U | X(s)\right).$$

We call such $X(t)$ a Markov process.

In other words, a stochastic process has the Markov property if its future evolution given all its history up to time s is the same as its future evolution given only its instantaneous position at time s.

Definition 1.9. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{M}$ a filtration. A stopping time is a function $\tau : \Omega \to [0, \infty)$ such that $\{\omega : \tau(\omega) \leq t\} \in \mathcal{M}_t$ for all $t \geq 0$.

In other words, we can think of a stopping time as a process where we can decide if $\tau \leq t$ at time t. For example, if we start flipping coins and set $\tau =$ {first time a heads comes up} then τ is a stopping time, since we will be able to tell whether our first heads comes up by the n^{th} flip or not (that is if $\tau \leq n$) by the time we flip the n^{th} coin. On the other hand if we set $\tau' =$ {the time of the first heads in the first time a sequence of 3 consecutive heads in a row} (that is, the start of the first time we get Heads, then Heads, then Heads again) then τ' is **not** a stopping time, since we cannot know in advance if the next two coins will be heads or not.

Definition 1.10. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{M}$ be a filtration and τ be a stopping time such that $\tau < \infty$ almost surely. A stochastic process $X(t) : \Omega \to \mathbb{R}^n$ that is adapted to $\mathcal{F}$ has the strong Markov property if for any $t > \tau$ and any open set $U \subset \mathbb{R}^n$ we have that

$$P\left(X(t) \in U | \mathcal{F}_\tau\right) = P\left(X(t) \in U | X(\tau)\right).$$

We call such $X(t)$ a strong Markov process.

Definition 1.11. Let $X(t)$ be a stochastic process taking values in $\mathbb{R}^n$ and $\mathcal{F}_t$ be a filtration. For $f : \mathbb{R}^n \to \mathbb{R}$ we say that $M(t) = f(X(t))$ is a martingale if for any s, t with $t > s$ we have

$$E\left[M(t)|\mathcal{F}_s\right] = M(s).$$

A martingale can be seen as a **fair game**, since we have that your expected gain/loss between times s and t, given by $E\left[M(t) - M(s)|\mathcal{F}_s\right]$ will be zero, since $E\left[M(t) - M(s)|\mathcal{F}_s\right] = E\left[M(t)|\mathcal{F}_s\right] - E\left[M(s)|\mathcal{F}_s\right] = M(s) - M(s) = 0$. A classic example of a martingale is the common betting on consecutive flips of a fair coin, where if the coin comes up heads you win \$1 and if it comes up tails you lose \$1. No matter how much money you've gained or lost by the s^{th} flip, your expected total if you keep playing until t flips given your money at time s is just what money you have at time s. This is because with each flip you don't expect to win or lose money. We now state an important theorem about martingales which is used later in the paper.

Theorem 1.12. ***(Doob's Martingale Inequality)*** *Let $M(t)$ be a martingale such that almost surely $t \mapsto M(t)$ is continuous. Then for any $p \geq 1$, any $\epsilon > 0$ and any $t \geq 0$ we have that*

$$P\left[\sup_{s\in[0,t]} |M(s)| \geq \epsilon\right] \leq \frac{1}{\epsilon^p} E\left[|M(t)|^p\right].$$

1.1.2 Brownian Motion

Although there are many different yet equivalent definitions of Brownian Motion, we use the one due to the French mathematician Paul Lévy. Before we do so, we introduce what is perhaps the most important distribution of all probability theory, the normal distribution.

Definition 1.13. The normal distribution with mean μ and variance σ^2, denoted $N(\mu, \sigma^2)$, is the probability distribution that has as its probability distribution function:

$$N(\mu, \sigma^2)[x] = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Which means that if $X : \Omega \to \mathbb{R}$ is a normally distributed random variable with mean μ and variance σ (denoted $X \sim N(\mu, \sigma^2)$), then for any $U \subset \mathbb{R}$ we have that

$$P(X \in U) = \int_U N(\mu, \sigma^2)[x]dx.$$

With this in mind, we now define the standard Brownian motion. This process is named after the Scottish botanist Robert Brown, who was one of the first people to describe it after observing it in the movement of pollen particles immersed in water in 1827 and publishing his findings in the paper [Bro28]. The process is also commonly called *Wiener process*, in honor of the US American mathematician Norbert Wiener, who dedicated much of his work to studying this subject.

Definition 1.14. A standard Brownian motion (also known as standard Wiener process) is a real-valued stochastic process $\{B(t) : t \geq 0\}$ such that:

- The process **starts at zero**: $B(0) = 0$.
- The process has **independent increments**: for any $t_0, t_1, \ldots, t_n$ such that $0 = t_0 \leq t_1 \leq \cdots \leq t_n$ the increments $B(t_i) - B(t_{i-1})$ are independent random variables for $i = 1, \ldots, n$.
- For all $t \geq 0$ and $h > 0$ the increment $B(t + h) - B(t)$ is **normally distributed** with mean 0 and variance h.
- Almost surely, the function $t \mapsto B(t)$ is **continuous**.

This definition by no means implies that there exist a stochastic process satisfying these properties. In fact, it is not trivial to prove that such a process exists. This is where the following theorem comes into play.

Theorem 1.15. ***(Wiener's Theorem)*** *The standard Brownian motion exists.*

The full proof of Wiener's theorem is standard and can be found in literature. The book [MP10] uses the same definition of Brownian motion as in this paper and contains this proof.

We now discuss some important properties of the standard Brownian motion. Namely, we will state the scaling invariance property, time inversion, nowhere differentiability of the standard Brownian motion, and most importantly, that the Brownian motion satisfies the strong Markov property as in Definition 1.10.

Theorem 1.16. ***(Scaling Invariance)*** *Let $B(t)$ be a standard Brownian motion and $a > 0$. Then the stochastic process $X(t) = \frac{1}{a}B\left(a^2 t\right)$ is another standard Brownian motion.*

Proof. We check that $X(t)$ satisfies all the requirements in Definition 1.14:

- $X(0) = \frac{1}{a}B(0) = 0$.
- $X(t_i) - X(t_{i-1}) = \frac{1}{a}\left[B(a^2 t_i) - B(a^2 t_{i-1})\right]$. Independence then follows from setting $t_i' = a^2 t_i$ and using the independent increment property for $B(t)$ and $0 = t_0' \leq t_1' \leq \cdots \leq t_n'$.
- $X(t+h) - X(t) = \frac{1}{a}\left[B(a^2 t + a^2 h) - B(a^2 t)\right]$. By the properties of $B(t)$ this random variable is normally distributed with mean 0 and variance $\frac{1}{a^2}(a^2 t + a^2 h - a^2 t) = \frac{1}{a^2}(a^2 h) = h$.
- Since $B(t)$ is almost surely continuous and the function g defined by $g(f(x)) = \frac{1}{a}f(a^2 x)$ is continuous if f is continuous, it follows that $X(t) = g(B(t))$ is almost surely continuous.

□

Theorem 1.17. ***(Time Inversion)*** *Let $B(t)$ be a standard Brownian motion. Then the stochastic process $X(t) = tB\left(\frac{1}{t}\right)$ is another standard Brownian motion.*

The proof of this fact is not too complicated and can be found in books such as [MP10].

Theorem 1.18. ***(Nowhere Differentiability)*** *The standard Brownian motion is almost surely nowhere differentiable. More specifically, for any $t \geq 0$ we have that:*

$$\text{either} \quad \limsup_{h \downarrow 0} \frac{B(t+h) - B(t)}{h} = \infty \quad \text{or} \quad \liminf_{h \downarrow 0} \frac{B(t+h) - B(t)}{h} = -\infty, \quad \text{or both.}$$

The two limits above are known respectively as upper and lower right derivatives of $B(t)$ and the full proof of this theorem can once again be obtained in [MP10].

Theorem 1.19. ***(Strong Markov Property)*** *The Brownian motion has the strong Markov property. More specifically, if $\{\mathcal{F}_t\}_{t \geq 0}$ is a filtration of the Brownian motion $B(t)$, we define $\mathcal{F}_t^+ = \cap_{s>t}\mathcal{F}_s$, and τ is a stopping time with $\tau < \infty$ almost surely, then*

$$\{B(\tau + t) - B(\tau) : t \geq 0\}$$

is another standard Brownian motion independent of $\mathcal{F}_\tau^+$.

Once again, the proof of the strong Markov property for Brownian motions can be found in almost any textbook on the subject, and this particular statement can be seen in [MP10].

2 The Dirichlet Problem

2.1 The Classical Dirichlet Problem and Heat Equilibrium

Suppose there is a metal rod of length l where each end is connected to a certain type of temperature control mechanism (for example, assume one end is connected to a fridge which maintains the constant temperature T_0 and the other sits at a small flame which maintains a constant temperature T_l). If these mechanisms maintain a constant temperature at each end of the rod, the temperature at the other points in the rod will change until the system reaches an equilibrium. More formally, if $T : [0, l] \to \mathbb{R}$ is the temperature function of the rod (that is, $T(x)$ is the temperature of the rod at point $x \in [0, l]$), then a natural question to ask is: after the system reaches equilibrium, what is the temperature at point x, that is, what is $T(x)$?

As it turns out, we can formulate this question as follows: find a twice continuously differentiable function (that is, a function in $C^2([0, l])$) $T : [0, l] \to \mathbb{R}$ such that:

- $\frac{d^2 T}{dx^2}(x) = 0$ for any $x \in (0, l)$.
- $T(0) = T_0$ and $T(l) = T_l$.

In the requirements above, the first one represents an intrinsic property of systems in temperature equilibrium which implies that the temperature at a given point is the average of its surrounding ones and the second one represents the fact that each end is maintained at a constant temperature established by outside factors.

This problem is not too hard to solve and has the linear solution

$$T(x) = \frac{(l-x)}{l}T_0 + \frac{x}{l}T_1 = T_0 + \frac{x}{l}(T_l - T_0).$$

Now suppose we have a similar problem, but instead of having a rod (which was taken to be a one-dimensional object) we have a metallic plate (a two-dimensional object) where every point of its boundary is kept at a certain temperature, or even a metal sculpture (a three-dimensional object) held under the same constraints. The generalization of this problem can in fact be made for any dimension and yields the classical Dirichlet problem.

Definition 2.1. (Classical Dirichlet Problem) Let $D \subset \mathbb{R}^n$ be a domain (meaning an open connected set) and $\phi : \partial D \to \mathbb{R}$ be a continuous function taking points from the boundary of D to the real numbers. The classical Dirichlet problem is to find a function $u \in C^2(D)$ satisfying:

i) $\Delta u(x) = 0$ for any $x \in D$, where Δ is the Laplace operator defined as $\Delta f(x) := \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(x)$.

ii) $\lim_{\substack{x \to y \\ x \in D}} u(x) = \phi(y)$ for all $y \in \partial D$.

As in the one dimensional case, the first requirement for the solution imposes an intrinsic property of heat equilibrium: the temperature at a point is the average of the temperature of the points surrounding it. This can be expressed as the *mean value property*, and translates to requiring that the Laplacian of the solution be zero everywhere inside D. A function satisfying this property is said to be *harmonic* in D. The second requirement relates the solution to the imposed temperatures at the boundary of D, forcing u to tend to the pre-determined constant temperatures when close to the boundary so that the solution can be seen as continuous on $D \cup \partial D$.

Although there exist ways of solving this problem that do not rely on probability theory, perhaps one of the most elegant is the one presented in Theorem 2.3. Before we state it, we must define the multidimensional version of Brownian motion.

Definition 2.2. Let $d \in \mathbb{N}$ be such that $d \geq 2$. We say the d-dimensional Brownian motion is the process $B(t) = \big(B_1(t), B_2(t), \dots, B_d(t)\big) \in \mathbb{R}^d$ where $B_1(t), B_2(t) \dots, B_d(t)$ are independent one-dimensional Brownian motions.

Theorem 2.3. *Consider we have the same setup as in Definition 2.1, let $B^x(t)$ denote a n-dimensional Brownian motion starting at $x \in D$ and set the stopping time $\tau = \inf\{t > 0 : B^x(t) \notin D\}$. Then if there is a solution to the classical Dirichlet problem, it is unique and equal to the function*

$$u(x) = E^x\left[\phi(\tau)\right],$$

where $E^x[\phi(\tau)]$ is the expected value of $\phi(\tau)$ for a Brownian motion starting at $X(0) = x \in D$.

We will discuss a generalized version of this approach soon, so the proof of this result is omitted.

2.1.1 A Note on "Bad" Domains

It should be noted that not all domains $D \subset \mathbb{R}^d$ have a solution to the Dirichlet problem. For example, let $d \geq 2$ and consider the domain $D := \{x \in \mathbb{R}^d : ||x|| < r^2\}\backslash\{0\}$ consisting of an open ball of radius r centered around the origin, with the origin deleted. If we define $F : \partial D \to \mathbb{R}$ by

$$F(x) = \begin{cases} 1 & \text{if } ||x|| = 1 \\ 0 & \text{if } x = 0 \end{cases}$$

there is no $f : D \to \mathbb{R}$ solving Dirichlet's problem. To see why, recall that for any $d \geq 2$ the $d-$dimensional Brownian motion is transient, meaning that $P^x\{F(X_\tau) = 0\} = 0$ for any $x \in D$. In fact, since F is constant on the set $\{x \in \mathbb{R}^d : ||x|| = r\}$ we get that $P^x\{F(X_\tau) = 1\} = 1$ and therefore any solution $f : D \to \mathbb{R}$ to the Dirichlet problem would have to satisfy $f(x) = 1$ for any $x \in D$. However, this implies

that $\lim_{x\to 0} f(x) = 1 \neq 0 = F(0)$, meaning that there does not exist a solution to the Dirichlet problem in this case.

As it turns out, we can find a solution to the problem as long as we only require the limit in Definition 2.1 to hold in what are called *regular* points of the boundary, as defined below.

Definition 2.4. Let $A \subset \mathbb{R}^d$ be a closed set, $B(t)$ be d-dimensional Brownian motion and τ_A be the stopping time $\tau_A = \inf\{t > 0 : B(t) \in A\}$. Then a point $x \in A$ is called regular if $P^x\{\tau_A = 0\} = 1$. Otherwise, the point is called irregular.

In the example above, the points $\{x : ||x|| = 1\}$ are regular for the closed set ∂D while the origin point $x = 0$ is not.

2.2 The Generalized Dirichlet Problem

As any good mathematician, once we know how to solve a problem a natural question to ask is: how can we make it *harder*? In order to obtain a harder and more general version of the Dirichlet problem we first introduce a very broad class of operators.

Definition 2.5. Let $D \subset \mathbb{R}^d$ be a domain. We say that an operator L on $C^2(D)$ is semi-elliptic if it is of the form

$$L = \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i} + \sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j},$$

where $b_i(x)$ and $a_{ij}(x)$ are continuous functions for all i, j and the matrix $a(x) := [a_{ij}(x)]_{i,j=1}^{d}$ is positive semi-definite (meaning that $a(x)$ only has non-negative eigenvalues at all $x \in D$) and symmetric (meaning that $a_{ij}(x) = a_{ji}(x)$ for all i, j).

Example 2.6. For example, the Laplacian operator Δ is semi-elliptic, since we have $b_i(x) = 0$ for any i and

$$a_{ij}(x) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases},$$

meaning that $a(x) = I$ is the identity matrix at all x. Since all the eigenvalues of I are 1 and all b_i, a_{ij} are continuous, Δ is semi-elliptic.

Semi-elliptic operators are of great importance in the study of Partial Differential Equations (PDEs) and motivate the generalized Dirichlet problem.

Definition 2.7. (Generalized Dirichlet Problem) Let $D \subset \mathbb{R}^n$ be a domain (meaning an open connected set), $\phi : \partial D \to \mathbb{R}$ be a continuous function taking points from the boundary of D to the real numbers and L be a semi-elliptic operator in D. The generalized Dirichlet problem is to find a function $u \in C^2(D)$ satisfying:

i) $Lu(x) = 0$ for any $x \in D$.

ii) $\lim_{x\to y, x\in D} u(x) = \phi(y)$ for all regular $y \in \partial D$.

As it turns out we can also use a probabilistic approach to solve this fully-deterministic problem! In order to do so, however, we will need to create generalizations of Brownian motions called *Itô diffusions*, and in order to do so, we will need to learn the basic tools of Stochastic Calculus. We will then show that the solution of the generalized Dirichlet problem can also be given as the expected value of $\phi(\tau)$, but we will need to tailor our stochastic process to work with a given semi-elliptic operator. That is, instead of using a Brownian motion $B(t)$ we will use an Itô diffusion $X(t)$ which is related to the operator L.

3 Basics of Stochastic Calculus

Itô Calculus was first presented to the world in the 1944 paper [Ito44], titled *Stochastic Integral* and written by the Japanese mathematician Kiyoshi Itô while he worked at the Nagoya Imperial University. The Itô integral has since found applications in a wide array of fields such as physics, finance and biology. In this section, we present some key concepts from Stochastic Calculus. In Section 3.1, we present the most important properties of the Itô integral, namely its construction and Itô's formula. In Section 3.2, we present Stochastic Differential Equations, focusing on some properties of Itô Processes that will help us tackle the generalized Dirichlet problem.

3.1 Itô Integrals

Our goal now is to establish what it means to integrate with respect to a Brownian motion. In other words, to define the expression

$$\int f(t,\omega)dB(t). \tag{1}$$

Intuitively, we can think of this as attempting to model the solution F to a differential equation of the form

$$\frac{d}{dt}F = f \cdot \text{“white noise”},$$

where the “white noise” term is random and has mean 0 and is taken to be the standard Brownian motion. Another possible interpretation of equation (1) is that of gambling: one can think of $B(t)$ as a random process you are betting on and $f(t,\omega)$ as the bet you are placing. The integral then represents the overall gain/loss over a certain period of time. We will also impose that $f(t,\omega)$ must be an *adapted process*, meaning it cannot see into the future, thus making the integral a fair game, that is, a *martingale*.

The integral in equation (1), known as Itô integral, can also be informally thought of as a Riemann-Stieltjes integral where the integrator is a standard Brownian motion, making $dB(t)$ a representation of the infinitesimal variation of $B(t)$. In order for the integral to be well defined, however, we must also require the integrand to be an adapted process (meaning that $f(t,\omega)$ cannot depend on events that happen after time t) and be in L^2.

3.1.1 Definition

Our construction of the Itô integral is similar to that of the Lebesgue integral, where we first define integration for simple functions and then extend this to a more general class of functions. For Itô's integral instead of using simple functions we use “elementary functions”. In order to define them, we first define the function space $\mathcal{V}$ of admissible integrands.

Definition 3.1. A function $\phi(t,\omega) : [0,\infty)\times\Omega \to \mathbb{R}$ is in $\mathcal{V}(S,T)$ if it satisfies the following three properties:

(i) $\phi(t,\omega)$ is $\mathcal{B}\times\mathcal{F}$-measurable.

(ii) For each $t \geq 0$, the function $\omega \mapsto \phi(t,\omega)$ is $\mathcal{F}_t$-measurable (this property is called being $\mathcal{F}_t$-adapted).

(iii) $E\left[\int_S^T f(t,\omega)^2 dt\right] < \infty$.

We now use this to help us define the elementary functions we are looking for.

Definition 3.2. A function $\phi(t,\omega) : [0,\infty)\times\Omega \to \mathbb{R}$ is elementary if $\phi \in \mathcal{V}$ and is of the form

$$\phi(t,\omega) = \sum_{j\geq 0} e_j \chi_{[t_t,t_{j+1}]}(t)$$

where e_j is constant for all j, and χ represents the indication function. That is, for any set $S \subset [0,\infty)$ we define $\chi_S(x) = \begin{cases} 1 \text{ if } x \in S \\ 0 \text{ if } x \notin S \end{cases}$.

We now define the Itô integral for elementary functions as:

Definition 3.3. (Itô Integral for Elementary Functions) If $\phi(t,\omega) : [0,\infty)\times\Omega \to \mathbb{R}$ is elementary, then we define

$$\int_S^T \phi(t,\omega)dB_t(\omega) := \sum_{j\geq 0} e_j \left(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)\right).$$

In other words, we sum over the values of e_j times the difference $\left(B_{t_{j+1}}(\omega) - B_{t_j}(\omega)\right)$. This is a random variable, and it has the property that $E\left[\int_S^T \phi(t,\omega)dB_t(\omega)\right] = 0$ for any ϕ. We can now expand this to a larger class of functions.

Definition 3.4. (Itô Integral) If $f(t,\omega) : [0,\infty)\times\Omega \to \mathbb{R}$ is in $\mathcal{V}(S,T)$, then we define

$$\int_S^T f(t,\omega)dB_t(\omega) := \lim_{n\to\infty}\int_S^T \phi_n(t,\omega)dB_t(\omega)$$

where $\{\phi_n\}_{n\geq 0}$ is a sequence of elementary functions with the property

$$\lim_{n\to\infty} E\left[\int_S^T \left(f(\omega,t)-\phi_n(\omega,t)\right)^2 dt\right] = 0.$$

It can be shown that such a sequence exists for any $f \in \mathcal{V}(S,T)$, but the proof of this fact is omitted in this paper. For a formal derivation of this fact, see [Øks03].

It should be noted that the usual rules for integration of deterministic integrals are not necessarily valid for stochastic ones. For example, we know that $\int_0^t s ds = \frac{1}{2}t^2$, while $\int_0^t B_s dB_s \neq \frac{1}{2}B_s^2$. In fact we will prove in the following section with the help of Itô's formula that $\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$.

3.1.2 Itô's Formula

Now that we have defined Itô's integral, which does not necessarily obey the usual laws of deterministic calculus, we must attempt to figure out what new rules can be used to help us study these objects. Among such rules, perhaps the most fundamental result of stochastic calculus, Itô's formula, requires us to first define a particular type of stochastic process called an *Itô process*.

Definition 3.5. Let $u, v \in L^2$ and X_0 be a starting point. An Itô Process is a stochastic process of the form $X_t = X_0 + \int_0^t u(s,B_s)ds + \int_0^t v(s,B_s)dB_s$. We also write this as $dX_t = udt + vdBt$.

Once we have this definition, we are ready to state Itô's Formula, which was first presented by Kiyoshi Itô in the 1951 paper [Ito51].

Theorem 3.6. ***(One Dimensional Itô Formula)*** *Let $dX_t = udt + vdBt$ be an Itô Process. Suppose $g \in C^2$ and $Y_t = g(t,X_t) \in \mathcal{L}_2$. Then Y_t is also an Itô Process and*

$$dY_t = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)dX_t^2. \tag{2}$$

Where $dX_t^2 = (udt + vdBt)^2$ and the following rules for multiplying differentials apply: $dt\cdot dt = 0$, $dt\cdot dB_t = dB_t\cdot dt = 0$ and $dB_t\cdot dB_t = dt$.

It should be noted that when expanded this formula may also be expressed as

$$dY_t = \left(\frac{\partial g}{\partial t}(t,X_t) + u\frac{\partial g}{\partial x}(t,X_t) + \frac{v^2}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\right)dt + \left(v\frac{\partial g}{\partial x}(t,X_t)\right)dB_t \tag{3}$$

Before we go on to explain why this theorem is true, we study the following results.

Lemma 3.7. *Let B_t be a standard Brownian motion. Then $E[B_t^4] = 3t^2$.*

Proof. We know that for $X \sim N(0,1)$ a normally distributed variable with mean 0 and variance 1, the fourth moment $E[X^4]$ is equal to 3. Additionally, by using the scaling property of Brownian motion from Theorem 1.16 we get that $\frac{B_t}{\sqrt{t}}$ has the same distribution as B_1, which is $N(0,1)$ and therefore

$$\begin{aligned} E[B_t^4] &= E\left[t^2\left(\frac{B_t}{\sqrt{t}}\right)^4\right] \\ &= t^2E[B_1^4] \\ &= 3t^2. \end{aligned}$$

□

Lemma 3.8. *Let $f : [0,\infty)\times\mathbb{R}\to\mathbb{R}$ be in L^2, $0 = t_0 < t_1 < \cdots < t_n = t$ be a partition of $(0,t)$, B_t be a standard Brownian motion and $\Delta B_i = B_{t_{i+1}} - B_{t_i}$. Then the following is almost surely true for the limit in L^2*

$$\lim_{\Delta t\to 0}\sum_i f(t_i, X_t)(\Delta B_i)^2 = \int_0^t f(s,x)ds.$$

Proof. First set $f_i = f(t_i, X_{t_i})$. To prove this lemma we recall that $f_i\left((\Delta B_i)^2 - \Delta t_i\right)$ and $f_j\left((\Delta B_j)^2 - \Delta t_j\right)$ will be independent whenever $i \neq j$ and then examine

$$\begin{aligned}
E\left[\left(\sum_i f_i(\Delta B_i)^2 - \sum_i f_i\Delta t\right)^2\right] &= E\left[\sum_{i,j} f_i\left((\Delta B_i)^2 - \Delta t_i\right) f_j\left((\Delta B_j)^2 - \Delta t_j\right)\right]\\
&= \sum_i E\left[f_i^2\left((\Delta B_i)^2 - \Delta t_i\right)^2\right]\\
&= \sum_i E\left[f_i^2\right] E\left[(\Delta B_i)^4 + (\Delta t_i)^2 - 2(\Delta B_i)^2\Delta t_i\right]\\
&= \sum_i E\left[f_i^2\right]\left(3(\Delta t)^2 + (\Delta t_i)^2 - 2(\Delta t_i)^2\right)\\
&= \sum_i E\left[f_i^2\right] 2(\Delta t_i)^2
\end{aligned}$$

Once we take the limit as $(\Delta t_i)\to 0$ the value above becomes 0. Since $\lim_{\Delta t\to 0}\sum_i f(t_i)\Delta t = \int_0^t f(s)ds$ we have thus proven the result for limits in L_2. □

The lemma above justifies the statement $(dB_t)^2 = dt$ and reveals to us one of the most important principles of Stochastic Calculus. Informally, this may be stated as "dB_t looks like a term of order $\frac{1}{2}$". Despite this being far from a rigorous statement, keeping it in mind can give a great deal of intuition for the proofs that lie ahead, especially when we study Itô-Taylor approximations.

Although the proof of Theorem 3.6 is not too hard once we have established Lemma 3.8, it will be omitted for the sake of conciseness, since it requires a non-trivial amount of computation. Having said that, the proof mostly consists of a Taylor expansion of g centered around X_t and evaluating the result as we take partitions where Δt become arbitrarily small.

In order to provide an example of applying Itô's formula we prove the following result.

Lemma 3.9. *For any $t \geq 0$ and any $n \in \mathbb{N}$ we have that $\int_0^t B_s^n dB_s = \frac{B_t^{n+1}}{n+1} - \frac{n}{2}\int_0^t B_t^{n-1}dt$.*

Proof. We begin by setting $X_t = B_t$ and taking $g(t,x) = \frac{x^{n+1}}{n+1}$, which gives us that $\frac{\partial g}{\partial t}(t,x) = 0$, $\frac{\partial g}{\partial x}(t,x) = x^n$ and $\frac{\partial^2 g}{\partial x^2}(t,x) = nx^{n-1}$. Now, using Itô's formula we get that:

$$d(g(t,B_t)) = \left(\frac{n}{2}B_t^{n-1}\right)dt + (B_t^n)dB_t$$

which then gives us that

$$\begin{aligned}
\frac{B_t^{n+1}}{n+1} &= g(0,B_0) + \int_0^t \frac{n}{2}B_t^{n-1}dt + \int_0^t B_t^n dB_t\\
&= \frac{n}{2}\int_0^t B_t^{n-1}dt + \int_0^t B_t^n dB_t
\end{aligned}$$

□

An analogous multidimensional version of what we have described so far can now be stated.

Definition 3.10. Let $u_i, v_{ij} : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}$ be in L_2 for all $i \in \{1,2,\ldots,d\}$ and $j \in \{1,2,\ldots,m\}$ and $B(t) = (B_1(t),\ldots,B_m(t))$ be a standard m-dimensional Brownian motion. A <u>d-dimensional Itô Process</u> $X(t) = (X_1,\ldots,X_d)$ is a process defined by d Itô processes of the form:

$$\begin{cases}
dX_1 = u_1 dt + v_{11}dB_1(t) + \cdots + v_{1m}dB_m(t)\\
\vdots\\
dX_n = u_n dt + v_{d1}dB_1(t) + \cdots + v_{dm}dB_m(t)
\end{cases}.$$

This may also expressed in matrix form

$$dX(t) = udt + vdB(t),$$

where $X(t)$ and $B(t)$ are defined as above and

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_d \end{bmatrix} \text{ and } v = \begin{bmatrix} v_{11} & \dots & v_{1m} \\ \vdots & \ddots & \vdots \\ v_{d1} & \dots & v_{dm} \end{bmatrix}.$$

Note that we do not require that $d = m$. This gives us more flexibility, since we can describe a d dimensional Itô process using a number of Brownian motions that is larger or smaller than d, depending on our specific need. We are now ready to state the more general, multidimensional version of Itô's formula.

Theorem 3.11. *(**Multidimensional Itô Formula**) Let $dX_t = udt + vdBt$ be an d-dimensional Itô Process. Suppose $g : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^q$ is C^2 and $Y_t = g(t, X(t)) \in L^2$. Then Y_t is a q-dimensional Itô Process and*

$$dY_k(t) = \frac{\partial g}{\partial t}(t, X(t))dt + \sum_i \frac{\partial g}{\partial x_i}(t, X(t))dX_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g}{\partial x_i \partial x_j}(t, X(t))dX_i dX_j.$$

Additionally, the following rules for multiplying differentials apply: $dt \cdot dt = 0$, $dt \cdot dB_i = dB_i \cdot dt = 0$ for all i and $dB_i \cdot dB_j = \delta_{ij}dt$, where $\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$ is the Kronecker delta.

3.2 Stochastic Differential Equations

Now that we have defined Itô processes (see Definition 3.10), we may start to think about how to explicitly write down a formula for them. Doing so comprises what we call *solving* the stochastic differential equation which defines the process. Below we provide an example of what solving a stochastic differential equation might look like.

Example 3.12. (Black-Scholes Model and the Geometric Brownian Motion) Consider the following stochastic differential equation derived from a technique in finance called the *Black-Scholes model*, where $\mu, \sigma \in \mathbb{R}$ are fixed. We wish to solve

$$dX(t) = \mu X(t)dt + \sigma X(t)dB(t). \tag{4}$$

In order to accomplish this, we will examine and try to find a closed formula for $\ln(X(t))$. By Itô's formula from equation (3) and setting $Y(t) = \ln(X(t))$ we get that

$$\begin{aligned} dY(t) &= \left(\mu X(t)\frac{1}{X(t)} - \frac{\sigma^2 X(t)^2}{2}\frac{1}{X(t)^2}\right)dt + \left(\sigma X(t)\frac{1}{X(t)}\right)dB(t) \\ &= \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dB(t), \end{aligned}$$

meaning that $Y(t)$ is just a Brownian motion with constant drift $\mu - \frac{\sigma^2}{2}$ and constant diffusion σ. That is, $Y(t) = Y(0) + \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)$, where $B(t)$ is a one-dimensional standard Brownian motion. Since we have that $Y(t) = \ln(X(t))$, we get that given the initial condition $X(0) = X_0$, the solution to (4) is

$$X(t) = X_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right],$$

a process commonly know as the *geometric Brownian motion*.

The example above, however, illustrates one of the few problems of this sort that we actually know how to solve. In most cases the solution cannot be reached analytically and we must instead resort to numerical methods as the ones which will be described in Section 4.

3.2.1 Itô Diffusions

In this section we define a subset of Itô processes which are guaranteed to have unique solutions for which the Markov property holds called *Itô diffusions* and lay out some of their important properties. We will then use these diffusions in Section 3.3 to solve the Dirichlet problem. We begin by defining these processes.

Definition 3.13. Let $X(t)$ be a d-dimensional Itô process defined by

$$dX(t) = b(X(t))\,dt + \sigma(X(t))\,dB(t), \tag{5}$$

where $b\colon \mathbb{R}^d \to \mathbb{R}^d$, $\sigma\colon \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are functions that do not explicitly depend on time and $B(t)$ is a standard m-dimensional Brownian motion. We say $X(t)$ is an Itô diffusion if there exists a constant D such that for all $x, y \in \mathbb{R}^n$

$$||b(x) - b(y)|| + ||\sigma(x) - \sigma(y)||_F \leq D||x - y||,$$

where $||\cdot||$ is the usual Euclidean norm in $\mathbb{R}^d$ and $||\cdot||_F$ is the Frobenius norm for matrices, defined by $||\sigma||_F = \sqrt{\sum_{i,j} |\sigma_{ij}|^2}$.

We will now state two very important properties of Itô diffusions. Namely, that (5) yields a unique solution and that this solution has the strong Markov property. We can intuitively understand why $X(t)$ must have the strong Markov property since in our definition we required the drift function b and the diffusion function σ to **not** depend on time explicitly, but rather only on the given position of the process $X(t)$, implying that the system has no memory of what happened in the past and that its evolution after any stopping time τ depends only on its given position $X(\tau)$ at time τ. The proof of the following theorems will be omitted, but the reader may refer to chapters 5 and 7, respectively, of [Øks03] for their complete derivation.

Theorem 3.14. *Let $X(t)$ be an Itô diffusion as in Definition 3.13. Then (5) has a unique t-continuous solution with the property that for any $T \in [0, \infty)$*

$$E\left[\int_0^T ||X(t)||^2 dt\right] < \infty.$$

Theorem 3.15. *Let $X(t)$ be an Itô diffusion as in Definition 3.13. Then $X(t)$ has the strong Markov property. That is, for any bounded Borel measurable function f on $\mathbb{R}^n$ and any stopping time τ such that $\tau < \infty$ almost surely we have that for any $h \geq 0$,*

$$E^x\left[f(X(\tau + h))\,|\mathcal{F}_\tau\right] = E^{X(\tau)}\left[f(X(h))\right].$$

Before we discuss numerical methods to solve SDEs, we define an important operator related to Itô diffusions that will be used in Section 5.

Definition 3.16. Let $X(t)$ be an Itô diffusion as in Definition 3.13, and $f\colon \mathbb{R}^n \to \mathbb{R}$ be a function. The infinitesimal generator of $X(t)$ is the operator $\mathcal{A}$ defined by

$$\mathcal{A}f(x) = \lim_{t\downarrow 0} \frac{E^x\left[f(X(t))\right] - f(x)}{t}. \tag{6}$$

Although we will not prove this fact, it can be shown that for an Itô diffusion of the form in (5) the infinitesimal operator will be well defined for any twice differentiable function with compact support (that is, any $f \in C_0^2(\mathbb{R}^n)$) and is be equal to

$$\mathcal{A}f(x) = \sum_{i=1}^d b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j}\left(\sigma\sigma^T\right)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x). \tag{7}$$

Example 3.17. Take the standard m-dimensional Brownian motion $B(t)$ defined by the equation $dB(t) = 0dt + I\,dB(t)$. That is, take $b(x) = 0$ and $\sigma(x) = I$ for any $x \in \mathbb{R}^n$. Then the infinitesimal operator of $B(t)$ is the Laplacian operator times $\frac{1}{2}$:

$$\mathcal{A}f(x) = \frac{1}{2}\sum_{i=1}^m \frac{\partial^2 f}{\partial x_i^2}(x) = \frac{1}{2}\Delta f(x).$$

3.3 Stochastic Dirichlet Problem and Proof of its Solution

Although we set out to solve the generalized Dirichlet problem as stated in definition 2.7, the proof of its solution requires knowledge about elliptic operators that is out of the scope of this paper. The solution in most cases, however, ends up being the same as a new version of the problem which is stated in terms of concepts from probability theory. Before we go on to describe this problem we define an important term.

Definition 3.18. Let $D \subset \mathbb{R}$ be a domain (meaning open and connected), $X(t)$ be an Itô diffusion on D and $f : D \to \mathbb{R}$ be a bounded, measurable function on D. We say f is X-harmonic if for any bounded open set U with $\overline{U} := U \cup \partial U \subset D$ we have that

$$f(x) = E^x \left[f(X_{\tau_U}) \right]$$

for any $x \in D$, where $\tau_U = \inf\{t > 0 : X(t) \notin U\}$ is the exit time of U.

Although seemingly disconnected from our original problem, the lemma below establishes why this is such an important concept: X-harmonic functions are ones for which the infinitesimal operator of X is zero in D.

Lemma 3.19. *Let $f : D \to \mathbb{R}$ be in $C^2(D)$. Then for any Itô diffusion $X(t)$ we have that*

$$f \text{ is } X\text{-harmonic in } D \iff \mathcal{A}_X f = 0 \text{ in } D.$$

We are now ready to state the stochastic version of the generalized Dirichlet problem.

Definition 3.20. Let $D \subset \mathbb{R}$ be a domain (meaning open and connected), $X(t)$ be an Itô diffusion on D, $\phi : \partial D \to \mathbb{R}$ be a bounded, measurable function and $\tau = \inf\{t > 0 : X(t) \notin D\}$ be the exit time of D. Then the stochastic Dirichlet problem is to find $u \in C^2(D)$ satisfying:

i) u is X-harmonic.

ii) $\lim_{t \uparrow \tau} u(X(t)) = \phi(X(\tau))$ almost surely.

Theorem 3.21. ***(Solution of the stochastic Dirichlet problem)*** *Consider the setup from Definition 3.20. Then there exists a unique solution to the the stochastic Dirichlet problem, which is given by*

$$u(x) = E^x \left[\phi(X(\tau)) \right].$$

Proof. (Sketch) Showing that u is indeed X-harmonic is quite intuitive, but requires us to define the harmonic measure for X and state the mean-value property, so this part will be omitted for the sake of conciseness. For a reference see chapter 7 of [Øks03].

The core of the proof relies on using a sequence of open sets $D_1, D_2, \ldots$ such that $D_i \subset D_j$ whenever $i < j$, $\overline{D_i} \subset D$ for all i and $\cup_{i=1}^{\infty} D_i = D$. Additionally we set $\tau_k = \inf\{t > 0 : X(t) \notin D_k\}$ for all k.

To show uniqueness, assume the bounded function $g : D \to \mathbb{R}$ solves the stochastic Dirichlet problem. We will prove that this implies that $g \equiv u$. To see why, notice that since g must be X-harmonic for any $x \in D$ we have that $g(x) = E^x \left[g\left(X(\tau_k) \right) \right]$ for any k. Additionally, since all functions are bounded we have that

$$g(x) = \lim_{k \to \infty} E^x \left[g\left(X(\tau_k) \right) \right] = E^x \left[\phi(X(\tau)) \right] = u(x),$$

thus proving uniqueness.

To show existence of the solution u we use the same sequence $\{D_i\}_{i \in \mathbb{N}}$ and notice that for any k, due to the strong Markov property for Itô diffusions,

$$\begin{aligned} u\left(X(\tau_k) \right) &= E^{X(\tau_k)} \left[\phi(X(\tau)) \right] \\ &= E^x \left[\phi(X(\tau)) \,\middle|\, \mathcal{F}_{\tau_k} \right]. \end{aligned}$$

If we set the value above as $M_k := E^x \left[\phi(X(\tau)) \,\middle|\, \mathcal{F}_{\tau_k} \right]$ for all k we see that the sequence $\{M_k\}_{k \in \mathbb{N}}$ forms a martingale. By a standard theorem about martingales (which can be found as Corollary C.9 in [Øks03]) we get that almost surely

$$\lim_{k \to \infty} M_k = \phi(X(\tau)). \tag{8}$$

Additionally, if we set $T = \max\{\tau_k, \min\{t, \tau_{k+1}\}\}$ we get that for each k the process $N_t = u(X(T)) - u\left(X(\tau_k)\right)$ is another martingale. Then by Doob's martingale inequality (Theorem 1.12), for any $\epsilon > 0$,

$$P\left[\sup_{t\in[\tau_k,\tau_{k+1}]}\left|u(X(t)) - u\left(X(\tau_k)\right)\right| > \epsilon\right] \leq \frac{1}{\epsilon^2}E^x\left[\left|u\left(X(\tau_{k+1})\right) - u\left(X(\tau_k)\right)\right|^2\right]. \tag{9}$$

Since for any $\epsilon > 0$ the right-hand side of the inequality goes to zero as k goes to infinity, equation (8) gives us that almost surely

$$\lim_{t\uparrow\tau} u(X(t)) = \phi(X(\tau)).$$

□

The last thing we do now is relate this problem to what was discussed in Section 2 by the following theorem.

Theorem 3.22. *Suppose we have the same setup as in Definition 2.7 and that L is uniformly elliptic in D (i.e.: the eigenvalues of $a(x)$ are bounded away from zero in D). If the Itô diffusion $X(t)$ has L as its infinitesimal generator, and $\tau = \inf\{t \geq 0 : X(t) \notin D\}$, then the function*

$$u(x) = E^x\left[\phi(X(\tau))\right]$$

is C^2 and solves the generalized Dirichlet problem from Definition 2.7.

The proof of this theorem is out of the scope of this paper, but can be found in section 9.2 of [Øks03]. In order to gain more intuition about how to implement such a solution, the reader may refer to the last section of this paper, where an example is carried out and computed numerically.

4 Numerical Methods for Stochastic Differential Equations

The goal of this section is to numerically simulate the solution of a stochastic differential equation of the form

$$dX(t) = a(t, X(t))dt + b(t, X(t))dB(t) \tag{10}$$

where $t \in [0, \infty)$, $X(t) \in \mathbb{R}^d$, $B(t)$ is a m-dimensional standard Brownian motion, $a : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is the drift coefficient and $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is the diffusion coefficient. In order to achieve this goal we first examine what it means for a given approximation to be a good one (meaning that it converges to what you want it to). After this has been established, we use a stochastic version of the deterministic Taylor's expansion in order to create discrete time approximation algorithms that can be used in a large array of applications.

Throughout this section we will use m to denote the dimension of the Brownian motion $B(t)$ and d to denote the dimension of the Itô Process in equation (10).

4.1 Discrete Time Approximations

Since the solution to most stochastic differential equations cannot be written out explicitly, we use computers to estimate what their approximate solution is at any given point in time. Computers, however, do not operate in continuous terms: there is a finite amount of memory and computational power available. This leads us to create discrete time approximation schemes, in which we start knowing what $X(t_0)$ is for some t_0 (or pick $X(t_0)$ from some probability distribution) and then use some method to determine an approximate probability distribution of $X(t)$ for $t > t_0$.

The general idea behind discrete time approximations is to split the difference $t - t_0$ in N equal parts $\Delta t = \frac{t-t_0}{N}$ and then create a sequence of N steps $Y_0, Y_1, \ldots, Y_N$ where $Y_0 = X(t_0)$ and Y_N has a similar probability distribution to that of $X(t)$. That is, we want that $Y_N \sim f(X(t)|X(t_0))$ where $f(X(t)|X(t_0))$ is the probability distribution function of $X(t)$ given that we know the value of $X(t_0)$.

In this sequence, Y_k will be an estimate of $X\left(t_0 + k\Delta t\right)$ and we will use this value to determine what Y_{k+1} is. In order words, in order to estimate $X(t)$ we progressively estimate $X(t_i)$ for several $t_i \in (t_0, t)$ satisfying $t_0 < t_1 < \cdots < t_N = t$.

4.1.1 Strong and Weak Convergence

Definition 4.1. Given the value of $X(t_0)$ and $t > t_0$, the discrete approximation Y_N of $X(t)$ converges strongly with order $\gamma > 0$ if there exists a constant $K < \infty$ such that for all step sizes $\Delta t = \frac{t-t_0}{N}$ we have that

$$E\left[\left|X_t - Y_N\right|\right] \leq K(\Delta t)^\gamma.$$

Strong approximations can give us a good way of running simulations of paths for Itô processes, but can become computationally expensive at times. For instance, the strong Euler-Maruyama method (which will be presented in Section 4.3) requires us to draw a normally distributed value at each iteration of the algorithm, which can be very costly if this will need to happen thousands of times for each simulation. The next definition presents us with an alternative approach, which despite not being as precise in terms of simulating sample paths gives us a much cheaper option which works on most applications.

Definition 4.2. Given the value of $X(t_0)$ and $t > t_0$, the discrete approximation Y_N of $X(t)$ converges weakly with order $\gamma > 0$ if for all polynomials g on $\mathbb{R}^d$ there exists a constant $K_g < \infty$ such that for all step sizes $\Delta t \in (0, 1)$ we have that

$$\left|E\left[g(X_t)\right] - E\left[g(Y_N)\right]\right| \leq K_g(\Delta t)^\gamma,$$

where $N \in \mathbb{N}$, $\Delta t = t/N$ and $X_0 = Y_0$.

As we shall see later, methods that give us weak convergence instead of strong convergence are generally computationally cheaper and are good enough for most applications that do not depend explicitly on the sample path but aim to approximate a function of that path. An example of this would be Monte Carlo estimations of the solution to the Dirichlet problem, where with each simulation we do not care about the specific trajectory but only on the exit location of the domain D (this will be explained in more detail in Section 5).

With weak approximations we often do not need to sample from normal distributions and may instead sample from discrete probabilities. For example, in the weak Euler-Maruyama method we need only to sample from the distribution

$$\Delta\tilde{B} = \begin{cases} \sqrt{\Delta t} & \text{with probability } \frac{1}{2} \\ -\sqrt{\Delta t} & \text{with probability } \frac{1}{2} \end{cases},$$

which is much cheaper than drawing from a normal distribution as computers tend to be more efficient at sampling from discrete distributions rather than complicated ones.

4.2 Itô-Taylor Expansions

The basis of most numerical methods used to solve deterministic differential equations relies on Taylor expanding a function, approximating some terms and then neglecting the remaining terms that are of higher order. When it comes to Itô processes, however, we are dealing with functions which are continuous everywhere but nowhere differentiable. This means that we cannot simply Taylor expand an Itô Process, since we cannot take derivatives at any point! The solution to this issue is to instead create a new version of Taylor expansions which is suitable for Itô processes. The outcome of this idea is the Itô-Taylor expansion, which relies on repeated applications of Itô's formula to the drift and diffusion coefficients. This method will then motivate the algorithms discussed in Section 4.3, such as a stochastic version of the deterministic Euler method.

4.2.1 Intuition for Single Variable Case

Definition 4.3. *(provisional)* Let $dX = a(X_t)dt + b(X_t)dB(t)$. Then we define the functionals L^0, L^1 as

$$L^0 f(x) = a(x)f'(x) + \frac{1}{2}b(x)^2 f''(x) \qquad \text{and} \qquad L^1 f(x) = b(x)f'(x)$$

where $f'(x) = \frac{df}{dx}(x)$ and $f''(x) = \frac{d^2 f}{dx^2}(x)$.

This allows us to write the univariate Itô's Formula from Theorem 3.6 in a nice way

$$f(X_t) = f(X_0) + \int_0^t L^0 f(X_s)ds + \int_0^t L^1 f(X_s)dB(s).$$

Since this equation is true for any suitable f we can even apply it to a and b to get

$$a(X_t) = a(X_0) + \int_0^t L^0 a(X_s)ds + \int_0^t L^1 a(X_s)dB(s)$$
$$b(X_t) = b(X_0) + \int_0^t L^0 b(X_s)ds + \int_0^t L^1 b(X_s)dB(s).$$

If we now apply these expressions to X_t we see that

$$\begin{aligned}
X_t &= X_{t_0} + \int_{t_0}^t a(X_s)ds + \int_{t_0}^t b(X_s)dB(s) \\
&= X_{t_0} + \int_{t_0}^t \left(a(X_{t_0}) + \int_{t_0}^s L^0 a(X_z)dz + \int_{t_0}^s L^1 a(X_z)dB(z) \right) ds \\
&\quad + \int_{t_0}^t \left(b(X_{t_0}) + \int_{t_0}^s L^0 b(X_z)dz + \int_{t_0}^s L^1 b(X_z)dB(z) \right) dB(s) \\
&= X_{t_0} + a(X_{t_0}) \int_{t_0}^t dt + \int_{t_0}^t \left(\int_{t_0}^s L^0 a(X_z)dz + \int_{t_0}^s L^1 a(X_z)dB(z) \right) ds \\
&\quad + b(X_{t_0}) \int_{t_0}^t dB(s) + \int_{t_0}^t \left(\int_{t_0}^s L^0 b(X_z)dz + \int_{t_0}^s L^1 b(X_z)dB(z) \right) dB(s) \\
&= X_{t_0} + a(X_{t_0})(t - t_0) + \int_{t_0}^t \left(\int_{t_0}^s L^0 a(X_z)dz + \int_{t_0}^s L^1 a(X_z)dB(z) \right) ds \\
&\quad + b(X_{t_0})(B(t) - B(t_0)) + \int_{t_0}^t \left(\int_{t_0}^s L^0 b(X_z)dz + \int_{t_0}^s L^1 b(X_z)dB(z) \right) dB(s).
\end{aligned}$$

If we set $\Delta t = (t - t_0)$, $\Delta B(t) = (B(t) - B(t_0))$ and $R(t) = \int_{t_0}^t \left(\int_{t_0}^s L^0 a(X_z)dz + \int_{t_0}^s L^1 a(X_z)dB(z) \right) ds + \int_{t_0}^t \left(\int_{t_0}^s L^0 b(X_z)dz + \int_{t_0}^s L^1 b(X_z)dB(z) \right) dB(s)$ we can write X_t in the more compact form

$$X_t = X_{t_0} + a(X_{t_0})\Delta t + b(X_{t_0})\Delta B(t) + R(t). \tag{11}$$

This provides the motivation behind the Euler-Maruyama method which will be presented in Section 4.3 and assumes the term R is small. We may also apply Itô's formula once again, this time to the integrand of the term $\int_{t_0}^t \int_{t_0}^s L^1 b(X_z)dB(z)dB(s)$, that is, to the function $L^1 b(X_z)$ to get that

$$\begin{aligned}
L^1 b(X_z) &= L^1 b(X_{t_0}) + \int_{t_0}^z L^0 L^1 b(X_u)du + \int_{t_0}^z L^1 L^1 b(X_u)dB(u) \\
&= b(X_{t_0})b'(X_{t_0}) + \int_{t_0}^z L^0 L^1 b(X_u)du + \int_{t_0}^z L^1 L^1 b(X_u)dB(u)
\end{aligned}$$

which tells us that

$$\begin{aligned}
\int_{t_0}^t \int_{t_0}^s L^1 b(X_z)dB(z)dB(s) &= \int_{t_0}^t \int_{t_0}^s \left(L^1 b(X_{t_0}) + \int_{t_0}^z L^0 L^1 b(X_u)du + \int_{t_0}^z L^1 L^1 b(X_u)dB(u) \right) dB(z)dB(s) \\
&= b(X_{t_0})b'(X_{t_0}) \int_{t_0}^t \int_{t_0}^s dB(z)dB(s) \\
&\quad + \int_{t_0}^t \int_{t_0}^s \left(L^1 b(X_{t_0}) + \int_{t_0}^z L^0 L^1 b(X_u)du + \int_{t_0}^z L^1 L^1 b(X_u)dB(u) \right) dB(z)dB(s)
\end{aligned}$$

In order to calculate the integral $\int_{t_0}^{t}\int_{t_0}^{s} dB(z)dB(s)$ we use Lemma 3.9.

$$\begin{aligned}\int_{t_0}^{t}\int_{t_0}^{s} dB(z)dB(s) &= \int_{t_0}^{t}\left[B(s)-B(t_0)\right]dB(s)\\ &= \frac{1}{2}\left(B(t)^2 - t - B(t_0)^2 + t_0\right) - B(t_0)(B(t)-B(t_0))\\ &= \frac{1}{2}\left[\left(B(t)-B(t_0)\right)^2 - (t-t_0)\right]\\ &= \frac{1}{2}\left((\Delta B(t))^2 - \Delta t\right).\end{aligned}$$

This motivates the Milstein method, by establishing the relation:

$$X_t = X_{t_0} + a(X_{t_0})\Delta t + b(X_{t_0})\Delta B(t) + b(X_{t_0})b'(X_{t_0})\frac{1}{2}\left((\Delta B(t))^2 - \Delta t\right) + R^*(t). \tag{12}$$

Where the new remainder $R^*(t)$ is equal to $R(t) - b(X_{t_0})b'(X_{t_0})\frac{1}{2}\left((\Delta B(t))^2 - \Delta t\right)$.

We may continue this process as much as we desire (as long as we can keep taking derivatives of a and b when needed), thus establishing a stochastic version of Taylor expansions. This motivates Theorem 4.11, which will be presented in Section 4.2.2 and gives us the intuition behind the algorithms presented in Section 4.3.

4.2.2 Multi-index, Other Useful Notation and Formal Statement for Multivariate Case

Before we state the multivariate version of Itô-Taylor expansions we must build some machinery in order to be able to express these expansions in an intelligible manner. Since these expansions can become very complicated as we venture into higher orders and higher dimensions, we dedicate most of this section to a notation system that allows us to state such Itô-Taylor expansions in a comprehensive way.

Definition 4.4. A multi-index of length $l \in \mathbb{N}$ is a row vector $\alpha = (a_1, \dots, a_l)$ with $a_1, \dots, a_l \in \{0, 1, \dots, m\}$ for some $m \in \mathbb{N}$. For any multi-index α we denote its length by $l(\alpha)$. We also define the multi-index of length zero as an element v such that $l(v) = 0$. We also set

$$\mathcal{M}(m) := \left\{\alpha = (a_1, \dots, a_l) : l \geq 0 \text{ and } a_1, \dots, a_l \in \{0, 1, \dots, m\}\right\}$$

as the set of all multi-indexes of any length with entries in $\{0, 1, \dots, m\}$.

Definition 4.5. Let $\alpha = (a_1, \dots, a_l)$ be a multi-index of length $l \geq 1$. Then we define $-\alpha$ and $\alpha-$ as $-\alpha := (a_2, \dots, a_l)$ and $\alpha- := (a_1, \dots, a_{l-1})$. That, is, $-\alpha$ deletes the first element of α and $\alpha-$ deletes the last. Both these new multi-indexes have length $l - 1$.

Definition 4.6. Let $\alpha = (a_1, \dots, a_l) \in \mathcal{M}(m)$ be a multi-index and ρ, τ two stopping times such that with probability one $0 \leq \rho \leq \tau \leq T$. Then for a function f we recursively define

$$I_\alpha[f(\cdot)]_{\rho,\tau} = \begin{cases} f(\tau) & \text{if } l(\alpha) = 0 \quad \text{(that is, if } \alpha = v\text{)}\\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho,s}ds & \text{if } l(\alpha) \geq 1 \text{ and } a_l = 0\\ \int_\rho^\tau I_{\alpha-}[f(\cdot)]_{\rho,s}dB_{a_l}(s) & \text{if } l(\alpha) \geq 1 \text{ and } a_l \neq 0\end{cases}$$

where $B_1, \dots, B_m$ are independent Brownian motions. We also use the notation $B_0(s) = s$ and $dB_0(s) = ds$. As a shorthand, when $f \equiv 1$ we also write

$$\begin{aligned}I_{\alpha,t} &:= I_\alpha[1]_{0,t}\\ &= \int_0^t\int_0^{s_l}\cdots\int_0^{s_2} dB_{a_1}(s_1)\dots dB_{a_{l-1}}(s_{l-1})dB_{a_l}(s_l)\end{aligned}$$

and when t is evident we may also omit it and just write I_α. Note that expression above may use the convention about B_0 mentioned.

Definition 4.7. Let $A \subset \mathcal{M}(m)$ be a set of multi-indexes. We say A is hierarchical if it has all the following properties:

- A is not the empty set: $A \neq \emptyset$.

- The the elements of A have bounded length: $\sup_{\alpha \in A} l(\alpha) < \infty$.
- For all $\alpha \in A \backslash \{v\}$ we have that $-\alpha \in A$.

Notice that this implies that all hierarchical sets must contain v, the element of zero length.

Definition 4.8. Let $A \subset \mathcal{M}(m)$ be a hierarchical set. Then we define the remainder set of A by:

$$B(A) := \{\alpha \in \mathcal{M}(m) \backslash A : -\alpha \in A\}.$$

We now revisit Definition 4.3 and substitute it for a more general version.

Definition 4.9. Let $a : [0, \infty) \times \mathbb{R}^d \to R^d$, $b^j : [0, \infty) \times \mathbb{R}^d \to R^d$ for all $j \in \{1, 2, \dots, m\}$ and $X(t) = X(t_0) + \int_{t_0}^t a(s, X(s))ds + \sum_{j=1}^m \int_{t_0}^t b^j(s, X(s))dB_j(s)$ be a d-dimensional Itô process. For any twice differentiable function $f : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ we define the functional L^0 as

$$L^0 := \frac{\partial}{\partial t} + \sum_{k=1}^{d} a^k \frac{\partial}{\partial x_k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b^{k,j} b^{l,j} \frac{\partial^2}{\partial x_k \partial x_l}, \tag{13}$$

where a^k represents the k^{th} component of a and $b^{k,j}$ the k^{th} component of b^j. Additionally, for any $j \in \{1, \dots, m\}$ we define the functional L^j as

$$L^j := \sum_{k=1}^{d} b^{k,j} \frac{\partial}{\partial x_k}.$$

This notation is defined this way in order to give us a more compact form of writing the multivariate Itô formula from Theorem 3.11:

$$dY_k(t) = L^0 g(s, X(s))ds + \sum_{j=1}^{m} L^j g(s, X(s))dB_j(s). \tag{14}$$

Definition 4.10. Let $\alpha = (a_1, \dots, a_l) \in \mathcal{M}(m)$ be a multi-index and $f : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ be a function with enough higher order partial derivatives. Then we recursively define

$$f_\alpha := \begin{cases} f & \text{if } l(\alpha) = 0 \quad \text{(that is, if } \alpha = v\text{)} \\ L^{a_1} f_{-\alpha} & \text{if } l(\alpha) \geq 1 \end{cases}.$$

In other words,

$$f_\alpha := L^{a_1} L^{a_2} \dots L^{a_l} f.$$

We now have all the notation required to state the multivariate Itô-Taylor expansion.

Theorem 4.11. *Let $a : [0, \infty) \times \mathbb{R}^d \to R^d$, $b^j : [0, \infty) \times \mathbb{R}^d \to R^d$ for all $j \in \{1, 2, \dots, m\}$ and*

$$X(t) = X(t_0) + \int_{t_0}^{t} a(s, X(s))ds + \sum_{j=1}^{m} \int_{t_0}^{t} b^j(s, X(s))dB_j(s)$$

be a d-dimensional Itô process with $t \in [t_0, T]$, ρ, τ be two stopping times such that with probability one $t_0 \leq \rho \leq \tau \leq T$, $A \subset \mathcal{M}(m)$ be a hierarchical set of multi-indexes and $f : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ be an appropriate function (meaning that it has enough higher order partial derivatives). Then

$$f(\tau, X(\tau)) = \sum_{\alpha \in A} I_\alpha[f_\alpha(\rho, X(\rho))]_{\rho,\tau} + \sum_{\alpha \in B(A)} I_\alpha[f_\alpha(\cdot, X(\cdot))]_{\rho,\tau}, \tag{15}$$

assuming a and b have enough high order partial derivatives.

The full proof of this theorem is in section 5.5 of [KP92].

Although seemingly complex when seen for the first time, the idea behind this theorem only requires you to repeatedly apply the multivariate Itô formula from Theorem 3.11 to the drift and diffusion coefficients of $X(t)$. Notice that if $f_\alpha(\rho, X(\rho))$ is a constant, we have that

$$\begin{aligned} I_\alpha[f_\alpha(\rho, X(\rho))]_{\rho,\tau} &= f_\alpha(\rho, X(\rho)) I_\alpha[1]_{\rho,\tau} \\ &= f_\alpha(\rho, X(\rho)) I_{\alpha,(\rho,\tau)}. \end{aligned}$$

Just as in the deterministic case, we can truncate the expression in equation (15) and ignore the remainder term $R_A(\rho,\tau) := \sum_{\alpha\in B(A)} I_\alpha[f_\alpha(\cdot, X(\cdot))]_{\rho,\tau}$ to obtain approximations of $X(\tau)$. Note that setting $\tau = t$ for some fixed $t \in [0,\infty)$ is also a stopping time. This gives rise to the Strong Itô-Taylor Approximations, which can be obtained for any hierarchical set A. The Milstein method for $m = 1$, presented in Section 4.3.2, for example, corresponds to the truncated expansion for the hierarchical set $A = \{v, (0), (1), (1,1)\}$.

There also exist the *Weak* Itô-Taylor Approximations in which the random variable $I_{\alpha,(\rho,\tau)}$ is not sampled directly from its distribution, but rather is approximated by another random variable $\widetilde{I}_{\alpha,(\rho,\tau)}$ which has the same moments as $I_{\alpha,(\rho,\tau)}$ up to some high enough order.

4.3 Algorithms

In this section we present algorithms that numerically approximate the solution to stochastic differential equations of the form in equation (10). We present both the strong and weak versions of these methods and how to implement them in an arbitrary number of dimensions. We first present the strong Euler-Maruyama method, which has strong order of convergence of $1/2$ and then its weak counterpart. We then present the Milstein method, which has strong order of convergence equal to 1.

4.3.1 Euler-Maruyama Method

The Euler-Maruyama method was developed by the Japanese mathematician Gisiro Maruyama as a generalization of the Euler method used to solve deterministic differential equations. For a given deterministic differential equation of the form

$$\frac{dy(t)}{dt} = f(t, y(t)),$$

the deterministic Euler method is given by

$$y_{n+1} = y_n + \Delta t f(t_n, y_n).$$

The justification behind this method is that it roughly approximates y_{n+1} using a first order Taylor expansion centered around y_n. The Euler-Maruyama method is a stochastic analogue, where we also take into account the Itô integral that appears in the right-hand side of equation (10).

The univariate Euler-Maruyama method (that is, when $d = m = 1$) is given by

$$Y_{n+1} = Y_n + a(t_n, Y_n)\Delta t + b(t_n, Y_n)\Delta B(t) \tag{16}$$

while in the multivariate case, the k^{th} component is given by

$$Y^k_{n+1} = Y^k_n + a(t_n, Y_n)\Delta t + \sum_{j=1}^{m} b^{k,j}(t_n, Y_n)\Delta B_k(t). \tag{17}$$

Due to the basic properties of Brownian motion, we know that $\Delta B(t) := B(t_{n+1}) - B(t_n)$ is a normally distributed variable with mean 0 and variance $t_{n+1} - t_n = \Delta t$, which means that we need not simulate a Brownian motion in order to make use of this algorithm; it is enough to draw $\Delta B(t)$ (or $\Delta B_k(t)$ in the multivariate case) from $N(0, \Delta t)$ at each iteration of the algorithm. Doing so guarantees strong convergence of order $\frac{1}{2}$ and, as we shall see in Section 4.3.2, in the case that the diffusion coefficient b does not depend on $X(t)$ (that is, it is constant or only a function of time) the algorithm actually has strong order of convergence equal to 1. This is because if $\frac{\partial b}{\partial X} = 0$, then the Milstein method reduces to Euler-Maruyama.

Given that drawing from normal distributions can be expensive when done thousands of times, we may instead use the *weak* Euler-Maruyama method, which is given by

$$Y^k_{n+1} = Y^k_n + a(t_n, Y_n)\Delta t + \sum_{j=1}^{m} b^{k,j}(t_n, Y_n)\Delta \widetilde{B}_k(t). \tag{18}$$

where $\Delta \widetilde{B}_k(t)$ is the random variable with mean 0 and variance Δt given by

$$\Delta \widetilde{B}_k(t) = \begin{cases} \sqrt{\Delta t} & \text{with probability } \frac{1}{2} \\ -\sqrt{\Delta t} & \text{with probability } \frac{1}{2} \end{cases}.$$

This method does not give strong convergence, but it does guarantee weak convergence as in Definition 4.2. The great benefit of this method is that it requires sampling from a uniform distribution, which is computationally a lot cheaper. Not only that, but in order to simulate $\Delta \widetilde{B}_k(t)$ we need only generate a single random bit! If that bit ends up being 1, we set $\Delta \widetilde{B}_k(t) = \sqrt{\Delta t}$ and if it is 0 we set $\Delta \widetilde{B}_k(t) = -\sqrt{\Delta t}$.

4.3.2 Milstein Method

As hinted with equation (12), the Milstein method for the univariate case (that is, when $d = m = 1$) is given by

$$Y_{n+1} = Y_n + a(t_n, Y_n)\Delta t + b(t_n, Y_n)\Delta B(t) + \frac{1}{2}b(t_n, Y_n)b'(t_n, Y_n)\left[(\Delta B(t))^2 - \Delta t\right]. \tag{19}$$

Or when written more compactly by omitting the arguments:

$$Y_{n+1} = Y_n + a\Delta t + b\Delta B + \frac{bb'}{2}\left[(\Delta B)^2 - \Delta t\right]. \tag{20}$$

If we have a multivariate Itô process (that is, $d > 1$) but maintain $m = 1$, then the Milstein method tells us to compute the k^{th} component by

$$Y_{n+1}^k = Y_n^k + a^k\Delta t + b^k\Delta B + \frac{1}{2}\sum_{l=1}^{d} b^l \frac{\partial b^k}{\partial x^l}\left[(\Delta B)^2 - \Delta t\right].$$

Once we venture into including multidimensional Brownian motions (that is, letting $m > 1$) we get the general and more complicated version:

$$Y_{n+1}^k = Y_n^k + a^k\Delta t + \sum_{j=1}^{m} b^{k,j}\Delta B_j + \sum_{j_1,j_2=1}^{m} L^{j_1} b^{k,j_2} I_{(j_1,j_2)} \tag{21}$$

where $I_{(j_1,j_2)}$ is taken from Definition 4.6.

The reason why this last version presented in equation (21) is more complicated is that we cannot write $I_{(j_1,j_2)}$ in simple terms when $j_1 \neq j_2$. If $j_1 = j_2$ then we calculated in Section 4.2.1 that $I_{(j_1,j_1)} = \frac{1}{2}\left[(\Delta B_{j_1})^2 - \Delta t\right]$, but when $j_1 \neq j_2$, $I_{(j_1,j_2)} = \int_{t_n}^{t_{n+1}}\int_{t_n}^{t_{n+1}} dB_{j_1} dB_{j_2}$ cannot be easily expressed using ΔB_{j_1}, ΔB_{j_2} and Δt, which limits the practicability of the Milstein method (and many other strong Itô-Taylor expansion methods) in higher dimensions. Having said that, some specific types of processes do not need to compute $I_{(j_1,j_2)}$ explicitly and make the Milstein method suitable for implementation.

5 Numerically Solving the Dirichlet Problem

In theorem 3.22 we use the Stochastic Calculus tools from Section 3 to tackle the Dirichlet problem. Doing so gives a solution which is given as the expectation of a random variable that is usually very hard or even impossible to solve analytically. Therefore, in this section we then use the numeric tools learned in section 4 allied with the Monte-Carlo Method to estimate the value of these solutions at some given points and illustrate the usefulness of the methods described in the paper so far.

5.1 Control: Testing For Accuracy and Other Factors

In order to test the efficiency of the method proposed we study one of the few problems for which there is an exact analytic solution. Consider the classical Dirichlet problem (that is, setting $L = \Delta$) where $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ is the open unit disk. If we use polar coordinates and set our boundary condition to be $\phi(\theta) = \sin(k\theta)$ for some $k \in \mathbb{N}$, then the exact solution for any $(r, \theta) \in D$ is $u(r, \theta) = r^k \sin(k\theta)$. To see why this is true, notice that u tends to the correct limit near the boundary and use the fact that in polar coordinates the Laplace operator becomes $\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}$, meaning that the Laplacian for $u(r, \theta) = r^k \sin(k\theta)$ is

$$\begin{aligned}
\Delta u(r,\theta) &= \frac{\partial^2}{\partial r^2}u(r,\theta) + \frac{1}{r}\frac{\partial}{\partial r}u(r,\theta) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}u(r,\theta) \\
&= k(k-1)r^{k-2}\sin(k\theta) + \frac{kr^{k-1}\sin(k\theta)}{r} + \frac{-k^2 r^k \sin(k\theta)}{r^2} \\
&= kr^{k-2}\sin(k\theta)\,(k - 1 + 1 - k) \\
&= 0.
\end{aligned}$$

For the sake of simplicity, we take $k = 2$, which makes our boundary function $\phi(\theta) = \sin(2\theta) = 2\cos(\theta)\sin(\theta) = 2xy$. The exact solution to this problem will be the function $u(r, \theta) = r^2 \sin(2\theta)$. The 3D plot of this solution forms a saddle, which can be seen in Figure 1.

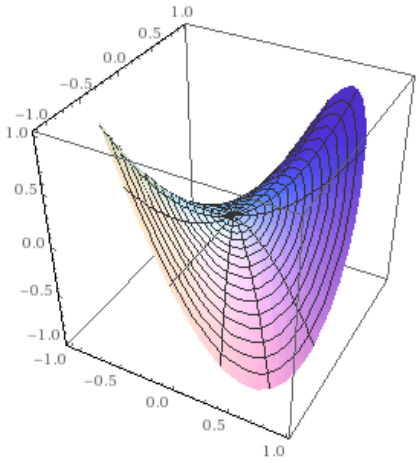

Figure 1: Exact solution $u(r, \theta) = r^2 \sin(2\theta)$. Graph obtained using WolframAlpha's "3D Parametric Plot" tool.

We then use a Monte-Carlo estimator to numerically obtain the solution at any given point within D. For more information on the Monte-Carlo method, the reader can refer to any standard textbook on the matter, or even section one of [FG17], a previous expository paper I wrote in this same program two years ago[2]. To summarize, in order to obtain an estimate of the expected value we wish to calculate we begin a weak Euler-Maruyama simulation of the diffusion process starting at the point we want to study and run it until the process leaves the domain D. Once this happens, we record the value of ϕ at the exit site and repeat the process several times (in most simulations we use N of the order of a few thousand simulations for each point). Our estimate then, is given by the average of these values, and its *statistical uncertainty* is the standard deviation of the data collected, divided by the square root of the number of simulations performed. That is, if $X_1, \dots, X_N$ are the values of ϕ at the exit point for our N trials, then the uncertainty δ is given by

$$\delta = \frac{\text{stdev}(X_1, \dots, X_N)}{\sqrt{N}}. \tag{22}$$

This δ accounts for the *statistical error* that comes from performing the Monte-Carlo method with a finite amount of data. By making N large, we can make this error as small as we would like. There is another source of error, however, called *systematic error* which is a result of our discretization of time. Again, in most cases by making Δt small, we can decrease it as much as we would like, but this error is much harder to estimate. We will present an experimental discussion later in this section which will attempt to establish what size for Δt is good enough.

It is important to stress that the uncertainty presented in equation (22) is by no means a strict error bound. As mentioned before, by making N large we can make this uncertainty become as small as we would like. This does not imply, however, that the system is infinitely precise, as the statistical uncertainty δ does not take into account the systematic error, which systematically shifts the results in a certain direction. To illustrate this, one can think of a person throwing darts at a target. If we make them throw a large number of darts we will get a good estimate of where they throw the dart on average, but this does not mean that this average will lie exactly on the target they aim at. There are systematic errors in play: the person might consistently throw darts below of the target due to lack of depth perception, or to the right of the target due to the particular choice of hand they use to throw the dart, etc. In our case, the discretization of time may make us over or underestimate the solution at any given point, something that cannot be measured by the statistical uncertainty δ.

The code used for these simulations was written in Python 3 and can be found in Appendix A. All data presented in the graphs were obtained using a personal computer and can be found in Appendix B. The error bars were calculated using the statistical error defined in equation (22). It should also be noted that once the particle leaves the domain, the value of ϕ is calculated at the first point outside the domain, rather than exactly at the boundary. As long as we make our step size small enough, continuity and boundedness assures us that since we will be close to the boundary this value will also be close enough to that of ϕ at a point in the boundary.

[2]This paper is available online at http://math.uchicago.edu/~may/REU2017/REUPapers/Guilhoto.pdf.

We first examine what happens when we fix an arbitrary angle $\theta = 1/2$ radians and vary the radius of points examined. With this θ, the exact solution is simply $r^2 \sin(2\theta) = r^2 \sin(1) \approx 0.84r^2$. We see the results in the graph below, where the solid line represents the exact solution and the dots represent the values collected at given points.

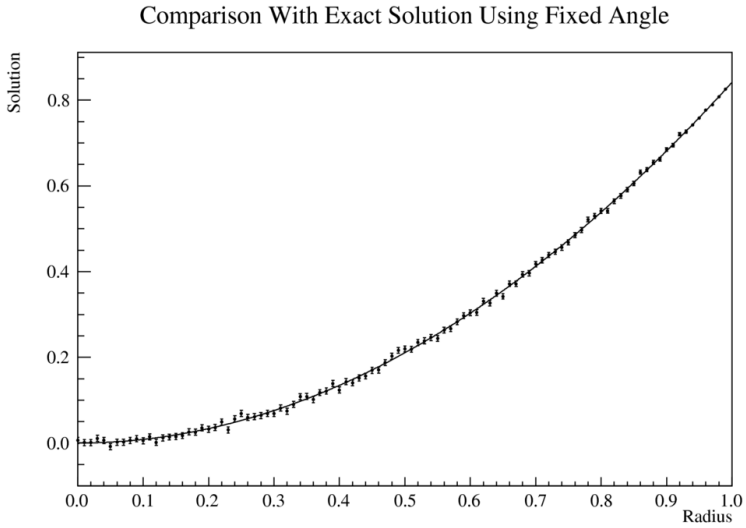

Figure 2: Results using a fixed angle of $\theta = 1/2$ and varying radii for the problem consisting of $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ and $\phi(\theta) = \sin(2\theta)$. The exact solution is shown as a solid line, while the numeric approximations are shown as points with their statistical error. Each point used a total of 10,000 trials and the time step was $\Delta t = 10^{-4}$.

We also examine what happens if we fix the radius at $r = 0.5$ and change the angle θ of points collected. In this case, the exact solution is $r^2 \sin(2\theta) = \sin(2\theta)/4$ and is once again presented as the solid line in the graph.

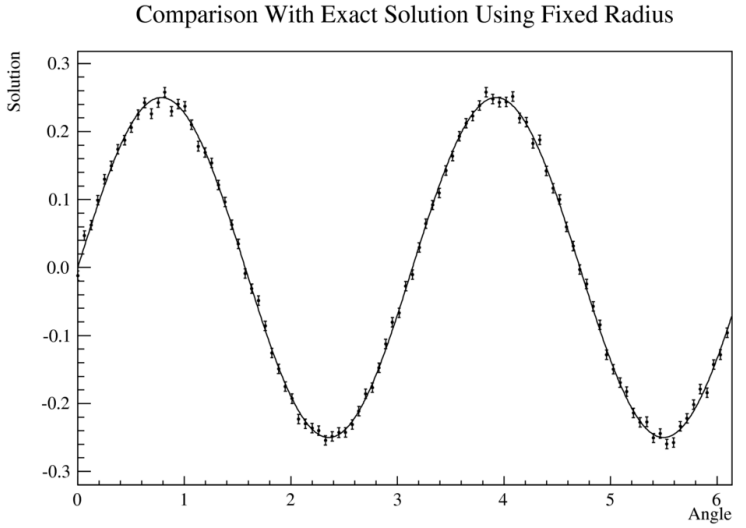

Figure 3: Results using a fixed radius of $r = 0.5$ and varying angles for the problem consisting of $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ and $\phi(\theta) = \sin(2\theta)$. The exact solution is shown as a solid line, while the numeric approximations are shown as points with their statistical error. Each point used a total of 10,000 trials and the time step was $\Delta t = 10^{-4}$.

As we can see, the method gives us great results! Furthermore, this experiment reveals to us some interesting properties of the Brownian motion. First of all, the reader might have noticed that in Figure 2 the statistical uncertainty of the points decrease as the radius increases. This is because once the process starts close to the boundary, the exit sites are more likely to fall within the same region. In comparison, when

$r = 0$ the probability of the process exiting at any given point in the unit circle is the same! This means that the standard deviation of the points collected will be greater when r is small, thus leading to a greater statistical uncertainty of the method for points farther away from the boundary.

Not only that, but it should also be noted that obtaining estimates for the solution close to the origin requires a much larger computational cost than for points close to the boundary. This is because as points are more distant from possible exit sites, on average they will require more steps before they leave D. This can indeed be observed in the graph below, where the average number of steps required for each simulation was recorded as a function of the radius. Note that the angle does not matter in this case, as it only changes the value of ϕ and we are now only concerned with what happens before the motion leaves D.

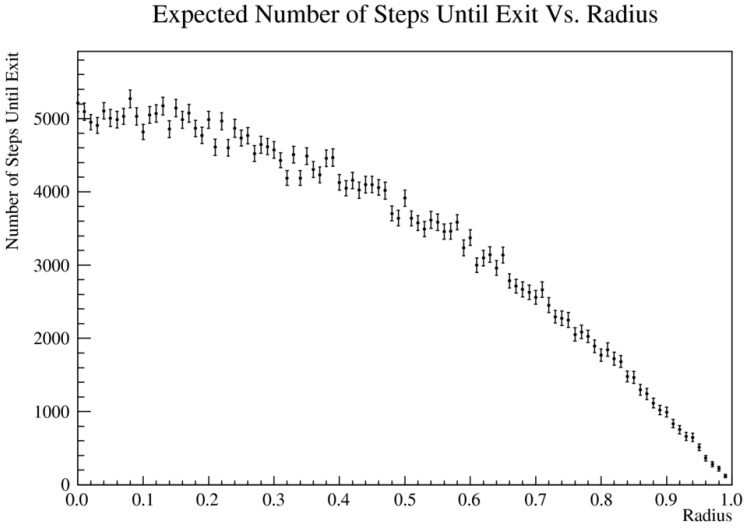

Figure 4: Experimental analysis of average number of steps required by the simulation until process leaves the unit circle D as a function of the radius. Numeric approximations for the averages are shown as points with their statistical error. Each point used a total of 1,000 trials and the time step was $\Delta t = 10^{-4}$.

Not only that, but the value of Δt we choose also has great influence in the average number of steps until exit. The graph below fixes the radius at $r = 0.5$ and varies the value of Δt in order to examine the relationship of that variable with number of steps until exit.

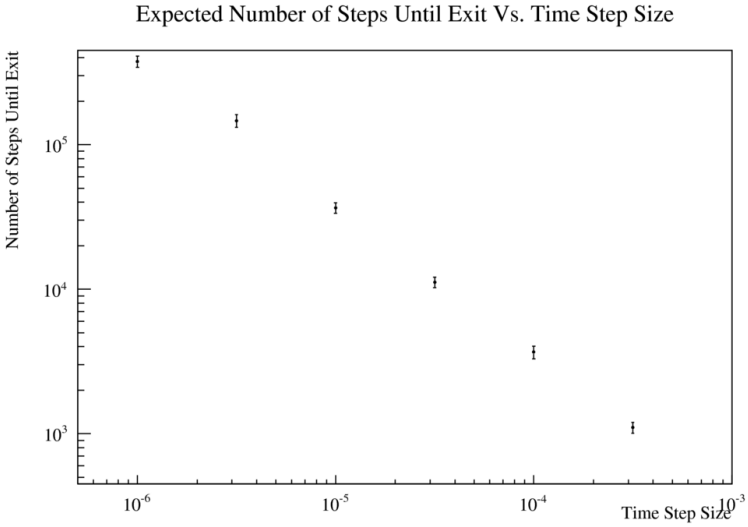

Figure 5: Experimental analysis of average number of steps required by the simulation until process leaves the unit circle D as a function of the time step size. Numeric approximations for the averages are shown as points with their statistical error. Each point used a total of 100 trials and the starting radius was $r = 0.5$. Note that both axes are displayed using logarithmic scale.

Since the bulk of computations for this method lies in the amount of steps required until exit, this can be used as a good estimate of the computational cost for trials and (usually) the amount of time required for computations. Since we now know the huge extent to which the time step size Δt affects computational cost a natural question to ask is: how small do we really need to make Δt in order to get good results? In reality we wish to pick a Δt which is small enough to give us the right answer, but not too small as to make the time of computation too big. The graph below explores this question, comparing the exact solution $u(0.5, \pi/4) = \frac{1}{4}$ with numerical results using different Δt.

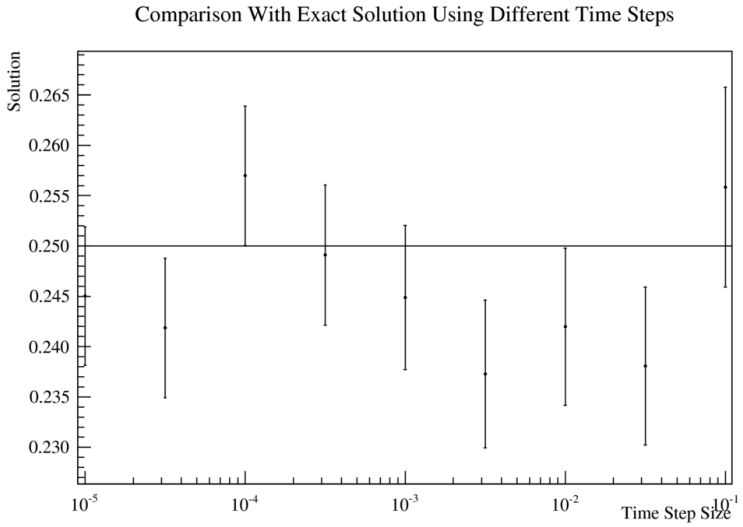

Figure 6: Experimental analysis of numeric result as a function of the time step size as compared with exact result 0.25 shown by the solid line. Numeric approximations for the averages are shown as points with their statistical error. Each point used a total of 10,000 trials, the starting radius was $r = 0.5$ and the starting angle was $\theta = \pi/4$. Note that the x-axis is presented in logarithmic scale.

As we can see, for this specific case we don't need to make Δt too small to get reasonable results. It should be said, however, that this particular problem employs a diffusion process that is homogeneous in space (that is, one for which the drift and diffusion coefficients are constant). Another problem for which that is not the case (such as the one in Section 5.2) might have a bigger dependence on Δt and thus require smaller step sizes and therefore a larger computational cost to attain good approximations.

5.2 Testing a More General Problem

Now that we have used the method to study a case for which we know the exact solution analytically, we move on to examine a more complicated one. To begin, for our domain we set the lemon shaped $D = D_1 \cup D_2$ where D_1 and D_2 are described as below. A picture of the boundary of the domain can be seen in Figure 7.

$$D_1 = \{(x, y) : xy \geq 0 \text{ and } |x|, |y| < 1\}, \quad D_2 = \{(x, y) : xy \leq 0 \text{ and } x^2 + y^2 < 1\} \tag{23}$$

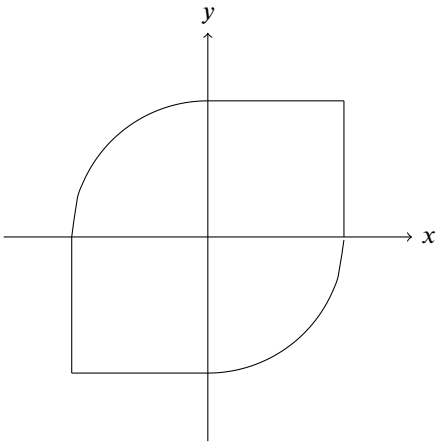

Figure 7: Complicated domain to be used in testing.

Additionally, we set ϕ to be the function $\phi(x, y) = \exp(\sin(xy))$ and define the operator L to be

$$L = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + 2.5\frac{\partial^2}{\partial x^2} + 12.5\frac{\partial^2}{\partial y^2} + 11\frac{\partial^2}{\partial x \partial y},$$

so that we use the Itô diffusion defined by

$$\begin{cases} dx = dX_1 = xdt + dB_1(t) + 2dB_2(t) \\ dy = dX_2 = ydt + 3dB_1(t) + 4dB_2(t) \end{cases}.$$

One can check that the diffusion matrix M of the Itô process $X(t)$ satistfies

$$\begin{aligned} \frac{1}{2}MM^T &= \frac{1}{2}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 2.5 & 5.5 \\ 5.5 & 12.5 \end{bmatrix}, \end{aligned}$$

which is the positive definite matrix which defines the second order derivative coefficients of L. It should also be noted that the eigenvalues of this matrix $\frac{1}{2}MM^T$ are bounded away from zero (they are $\frac{15\pm\sqrt{221}}{2} > 0$), meaning that L really is a uniformly elliptic operator and therefore that we may take into consideration Theorem 3.22. Additionally, since the diffusion matrix M is constant, the Milstein method reduces to the the Euler-Maruyama method, giving us a higher order of convergence for the results.

Although I am unable to prove that this problem does not have an analytical solution, it is definitely beyond my reach to obtain one, even if it exits. Having said that, it is similar enough in shape to the case tested in the previous section (with just some added ugliness), so we should expect a solution that is also somewhat similar in shape. Therefore, we employ the numerical methods discussed in this paper to obtain approximations of the solution at any given point.

We begin by observing points using the fixed radius of $r = 0.5$. As expected, we see in Figure 8 that we a obtain solution that is similar in shape to that of the problem in Section 5.1. It should be said, however, that although the shape of the solution resembles a sine wave of sorts, attempting to perform a fit of the form $\alpha\sin(\beta\theta) + \gamma$ with α, β and γ as free variables does not give us a nice fit, meaning that the solution is not as simple as the one we had in the last problem and that this numerical method is indeed useful.

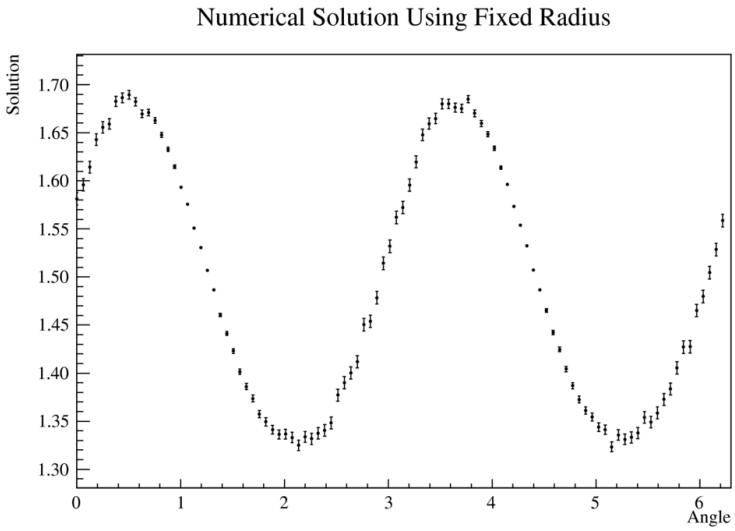

Figure 8: Results using a fixed radius of $r = 0.5$ and varying angles for the problem described in the beginning of Section 5.2. The numeric approximations are shown as points with their statistical error. Each point used a total of 10,000 trials and the time step was $\Delta t = 10^{-5}$.

We then examine what happens in the fourth quadrant once we fix the angle $\theta = 7\pi/4$ and vary the radii. Once again we obtain a solution which is similar in shape to the one presented in Figure 2 (with the sign flipped). It should be stressed again, however, that this solution does not agree with the fit of a parabola. If the reader looks at the behaviour of the solution once r gets close to 1 we see that it stops decreasing as much, leveling a bit and suggesting the existence of an inflection point somewhere between $r = 0.4$ and $r = 0.9$.

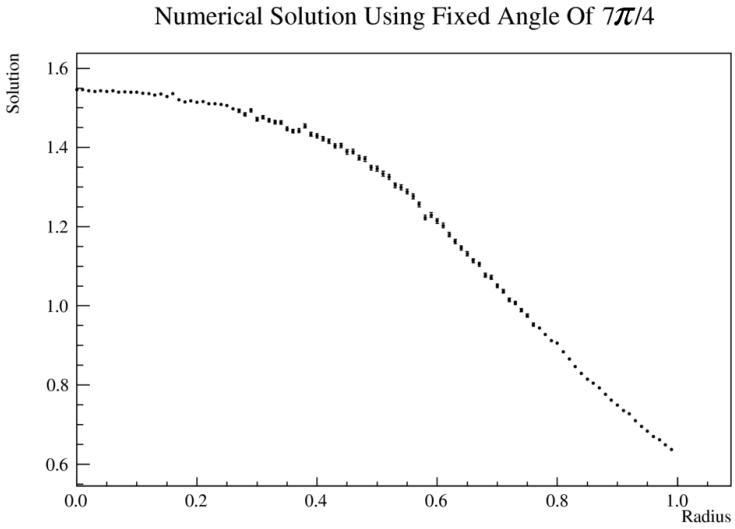

Figure 9: Results using a fixed angle of $\theta = 7\pi/4$ and varying radii for the problem described in the beginning of Section 5.2. The numeric approximations are shown as points with their statistical error. Each point used a total of 10,000 trials and the time step was $\Delta t = 10^{-5}$.

If we look at the first quadrant and this time fix $\theta = \pi/4$ we see that the solution once again resembles what we saw in the previous problem. Note that in this case we may pick any $r \in [0, \sqrt{2})$ and our point will still be inside the domain D.

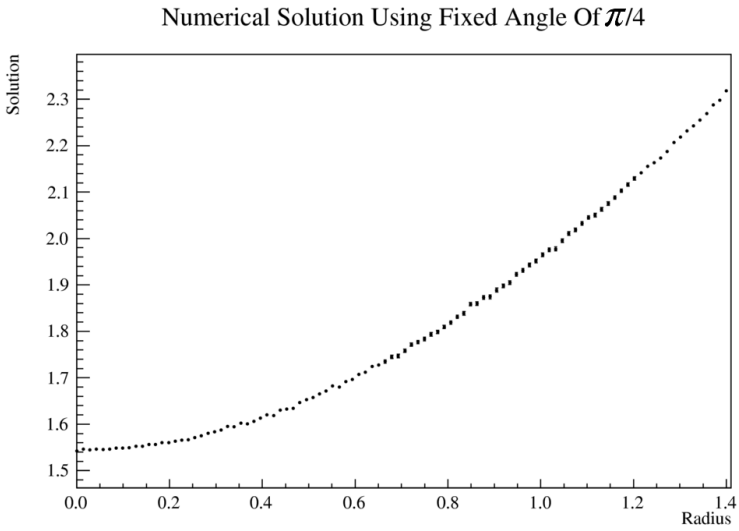

Figure 10: Results using a fixed angle of $\theta = \pi/4$ and varying radii for the problem described in the beginning of Section 5.2. The numeric approximations are shown as points with their statistical error. Each point used a total of 10,000 trials and the time step was $\Delta t = 10^{-5}$.

It is important to stress that since in this case the drift coefficient depends on the location of the process at any given step, the systematic error has a larger dependence on Δt. If we compare results obtained when Δt is 10^{-4} versus when it is 10^{-6} (presented in Figure 11) we can see that although the results are not entirely unrelated, there is a noticeable difference between them. More specifically, we can see that almost every point collected using $\Delta t = 10^{-6}$ is considerably lower than their counterpart using $\Delta t = 10^{-4}$. The odds of this happening due to sheer luck (that is, caused by the statistical error) is very slim, leading us to conclude that the disparity comes from a systematic error due to time discretization.

A possible explanation for why the estimates using $\Delta t = 10^{-4}$ were larger than what was obtained for $\Delta t = 10^{-6}$ is that the code used for the simulations calculates the value of ϕ at the first point obtained *outside* the domain D. Since $\phi(x, y) = \exp(\sin(xy))$ will be larger for such points when compared to the boundary close to them, it makes sense that using a larger step size will result in overestimates, as they are farther away from the boundary and therefore have a larger value for ϕ. Once we decrease the time step size to $\Delta t = 10^{-6}$, the first points obtained outside of D are much closer to the boundary, thus decreasing the systematic error.

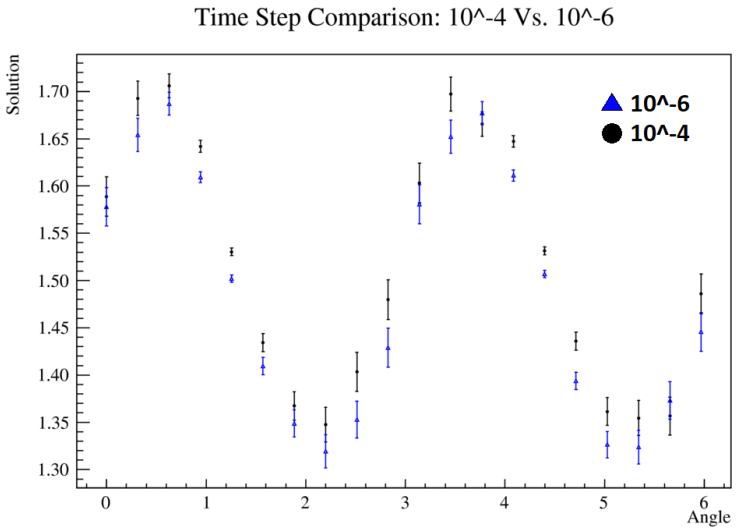

Figure 11: Experimental analysis of numeric result as a function of the time step size for the problem described in the beginning of Section 5.2. Results using $\Delta t = 10^{-4}$ are represented by black circles and ones using $\Delta t = 10^{-6}$ are represented by blue triangles. Numeric approximations for the averages are shown as points with their statistical error. Each point used a total of 1,000 trials and the starting radius was $r = 0.5$.

From this case and its comparison with the problem presented in Section 5.1 we may conclude that the more non-homogeneous our drift and diffusion coefficients are (that is, the more they vary according to time and space), the smaller we need to make our time step size Δt in order to make the approximations obtained using the algorithms described in Section 4.3 realistic. This is also the case when the drift and diffusion are larger, since a large step size might result in ϕ being calculated farther from the boundary and thus leading to systematic errors such as the ones seen in Figure 11.

Acknowledgements

This paper was only made possible due to the University of Chicago's Research Experience for Undergraduates. This program has had a profound impact in my undergraduate career and I could not be more thankful to all the people involved in its organization, especially professor Peter May, who year after year puts in an unbelievable amount of work into it and assures international students like me are able to obtain funding for the Summer. I would also like to deeply thank my mentor Jon DeWitt for his continued support, great feedback and willingness to learn about a new topic with me.

References

[Bro28] Robert Brown. A brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

[FG17] Leonardo Ferreira Guilhoto. Applying markov chains to monte carlo integration. *The University of Chicago Research Experience for Undergraduates*, 2017.

[Ito44] Kiyosi Ito. Stochastic integral. *Proceedings of the Imperial Academy*, 1944.

[Ito51] Kiyosi Ito. On a formula concerning stochastic differentials. *Nagoya Math. J.*, 3:55–65, 1951.

[KP92] Peter E. Kloeden and Eckhard Platen. *Numerical Solution to Stochastic Differential Equations*. Springer-Verlag, 1992.

[LL10] Gregory F. Lawler and Vlada Limic. *Random Walk: A Modern Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[MP10] Peter Mörters and Yuval Peres. *Brownian Motion*. Cambridge University Press, 2010.

[Øks03] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2003.

Appendix A: The Code

For the sake of completeness, the original code used in the implementations seen in Section 5 is presented in this Appendix. Everything was programmed using Python 3 and ran on a personal computer. No advanced features of Python were used as to make the code more readable to non-programmers. The code for the control case in Section 5.1 and the more complicated case in problem 5.2 are presented separately for the same reason. Below is the code used in Section 5.1, and the code used in Section 5.2 is presented immediately after.

The reader should feel free to use the code here for any non-profit oriented end they wish, but do give appropriate credit if that is the case. I do not hold myself responsible for any unintended result of third parties using this code.

```
### Libraries
import random
import numpy as np
import statistics


def phi(x,y):
    '''
    Phi is the boundary codition, which is defined as:
        phi(theta)=sin(2*theta).
    Since we are only looking at the unit disk we get that:
        phi(x,y)=2*sin(theta)*cos(theta)=2*x*y.

    Input: Euclidean coordinates (x,y) in the unit disk.
    Output: value of Phi at that point.
    '''
    return 2*x*y

def weak_euler_maruyama_step(var,sqrt_delta_t):
    '''
    Performs one step of the weak Euler-Maruyama method for a Brownian
    motion. This function only performs a one-dimensional step.

    Input: var is an Euclidean coordinate; sqrt_delta_t is the square
           root of the time step.
    Output: New Euclidean coordinate after an iteration of the
            Euler-Maruyama algorithm.
    '''
    rand_bit = random.randint(0,1)
    if rand_bit == 1:
        return var + sqrt_delta_t
    else:
        return var - sqrt_delta_t

def in_range(x,y):
    '''
    Checks if (x,y) is in the open unit disk.

    Input: Euclidean coordinates of a point in the plane.
    Output: True if the point is within the open unit disk, False
            otherwise.
    '''
    r_sqrd = x**2 + y**2
    if r_sqrd < 1:
        return True
    else:
        return False
```

```
def simulate_trial(x_0,y_0,delta_t):
    '''
    Simulates a trial of the Monte Carlo estimator. That is, uses the
    weak Euler-Maruyama method to simulate a Brownian motion until it
    reaches the boundary of D.

    Input: starting position (x_0,y_0) of the Brownian motion and the
           time step delta t.
    Output: the value of phi at the exit point of that simulation.
    '''
    x = x_0
    y = y_0
    sqrt_delta_t = np.sqrt(delta_t)
    while(in_range(x,y)):
        x = weak_euler_maruyama_step(x,sqrt_delta_t)
        y = weak_euler_maruyama_step(y,sqrt_delta_t)
    return phi(x,y)


def monte_carlo(x_0,y_0,delta_t,n):
    '''
    Performs a Monte-Carlo estimate of u(x_0,y_0) using n simulations.

    Input: Euclidean coordinates (x_0,y_0) of the point to be tested;
           the time step Delta t; and the number n of simulations to be
           carried out.
    Output: The pair [mean, uncertainty] of the Monte-Carlo estimate of
            the solution at the point (x_0,y_0).
    '''
    random.seed()
    results = []
    for i in range(n):
        results.append(simulate_trial(x_0,y_0,delta_t))
    mean = statistics.mean(results)
    uncertainty = statistics.stdev(results)/np.sqrt(n)
    return [mean,uncertainty]
```

Finally, below is the original code used in Section 5.2.

```
#### Libraries
import random
import numpy as np
import statistics

def phi(x,y):
    '''
    Phi is the boundary codition, which is defined as phi(x,y)=e^sin(xy).

    Input: Euclidean coordinates (x,y) in boundary of D.
    Output: value of Phi at that point.
    '''
    return np.exp(np.sin(x*y))

def instant_drift(position):
    '''
    Input: position is a pair corresponding to Euclidean coordinates [x,y].
    Output: The drift vector of the Ito diffusion at that point.
    '''
```

```
    return position

def instant_diffusion(position):
    '''
    Input: position is a pair corresponding to Euclidean coordinates [x,y].
    Output: The diffusion matrix of the Ito diffusion at that point.
    '''
    return [[1,2] , [3,4]]

def weak_euler_maruyama_step(position,delta_t,sqrt_delta_t):
    '''
    Performs one step of the weak Euler-Maruyama method for an Ito
    diffusion. This function performs a two-dimensional step.

    Input: position is a pair corresponting to Euclidean coordinates;
           delta_t is the time step;
           sqrt_delta_t is the square root of the time step.
    Output: New Euclidean position after an iteration of the
            Euler-Maruyama algorithm.
    '''
    drift = instant_drift(position)
    diffusion = instant_diffusion(position)
    x = position[0] + drift[0]*delta_t
    y = position[1] + drift[1]*delta_t
    rand_num = random.randint(0,3)
    if rand_num == 0:
        x = x - (diffusion[0][0] + diffusion[0][1])*sqrt_delta_t
        y = y - (diffusion[1][0] + diffusion[1][1])*sqrt_delta_t
    elif rand_num == 1:
        x = x + (-diffusion[0][0] + diffusion[0][1])*sqrt_delta_t
        y = y + (-diffusion[1][0] + diffusion[1][1])*sqrt_delta_t
    elif rand_num == 2:
        x = x + (diffusion[0][0] - diffusion[0][1])*sqrt_delta_t
        y = y + (diffusion[1][0] - diffusion[1][1])*sqrt_delta_t
    else:
        x = x + (diffusion[0][0] + diffusion[0][1])*sqrt_delta_t
        y = y + (diffusion[1][0] + diffusion[1][1])*sqrt_delta_t
    return [x,y]

def in_range(position):
    '''
    Checks if (x,y) is in the domain D.

    Input: Euclidean coordinates of a point in the plane.
    Output: True if the point is within D, False otherwise
    '''
    if (position[0]**2+position[1]**2 < 1):
        return True
    elif (position[0]*position[1] < 0):
        return False
    elif (abs(position[0])<1) and (abs(position[1])<1):
        return True
    else:
        return False

def simulate_trial(x_0,y_0,delta_t,sqrt_delta_t):
    '''
    Simulates a trial of the Monte Carlo estimator. That is, uses the
```

```
    weak Euler-Maruyama method to simulate a Brownian motion until it
    reaches the boundary of D.

    Input: starting position x_0,y_0 of the Brownian motion;
           the time step delta t;
           the square root of the time step delta t.
    Output: the value of phi at the exit point of that simulation.
    '''
    position = [x_0,y_0]
    while(in_range(position)):
        position = weak_euler_maruyama_step(position,delta_t,sqrt_delta_t)
    return phi(position[0],position[1])

def monte_carlo(x_0,y_0,delta_t,n):
    '''
    Performs a Monte-Carlo estimate of u(x_0,y_0) using n simulations.

    Input: Euclidean coordinates (x_0,y_0) of the point to be tested;
           time step Delta t;
           the number n of simulations to be carried out.
    Output: The pair [mean, uncertainty] of the Monte-Carlo estimate of
            the solution at the point (x_0,y_0).
    '''
    random.seed()
    sqrt_delta_t = np.sqrt(delta_t)
    results = []
    for i in range(n):
        results.append(simulate_trial(x_0,y_0,delta_t,sqrt_delta_t))
    mean = statistics.mean(results)
    uncertainty = statistics.stdev(results)/np.sqrt(n)
    return [mean,uncertainty]
```

Appendix B: Data

Since the algorithms implemented in Section 5 are random in nature, the data for the graphs in that section are presented in this appendix for the sake of completeness.

Table 1: Data for Figure 2.

r	Sol.	δ	r	Sol.	δ	r	Sol.	δ	r	Sol.	δ
0	0.006	0.007	0.25	0.069	0.007	0.5	0.219	0.007	0.75	0.468	0.006
0.01	0.001	0.007	0.26	0.060	0.007	0.51	0.219	0.007	0.76	0.485	0.006
0.02	0.001	0.007	0.27	0.061	0.007	0.52	0.234	0.007	0.77	0.497	0.006
0.03	0.011	0.007	0.28	0.064	0.007	0.53	0.238	0.007	0.78	0.522	0.006
0.04	0.006	0.007	0.29	0.069	0.007	0.54	0.246	0.007	0.79	0.530	0.006
0.05	-0.008	0.007	0.3	0.069	0.007	0.55	0.244	0.007	0.8	0.542	0.005
0.06	0.002	0.007	0.31	0.082	0.007	0.56	0.263	0.007	0.81	0.542	0.006
0.07	0.002	0.007	0.32	0.074	0.007	0.57	0.266	0.007	0.82	0.564	0.005
0.08	0.006	0.007	0.33	0.090	0.007	0.58	0.283	0.007	0.83	0.577	0.005
0.09	0.010	0.007	0.34	0.108	0.007	0.59	0.297	0.007	0.84	0.591	0.005
0.1	0.005	0.007	0.35	0.108	0.007	0.6	0.304	0.007	0.85	0.606	0.005
0.11	0.015	0.007	0.36	0.101	0.007	0.61	0.305	0.007	0.86	0.632	0.005
0.12	0.002	0.007	0.37	0.118	0.007	0.62	0.331	0.007	0.87	0.638	0.005
0.13	0.012	0.007	0.38	0.121	0.007	0.63	0.327	0.007	0.88	0.655	0.005
0.14	0.014	0.007	0.39	0.139	0.007	0.64	0.349	0.007	0.89	0.662	0.005
0.15	0.015	0.007	0.4	0.124	0.007	0.65	0.342	0.006	0.9	0.685	0.004
0.16	0.017	0.007	0.41	0.143	0.007	0.66	0.372	0.006	0.91	0.695	0.004
0.17	0.026	0.007	0.42	0.140	0.007	0.67	0.372	0.006	0.92	0.720	0.004
0.18	0.025	0.007	0.43	0.151	0.007	0.68	0.394	0.006	0.93	0.726	0.004
0.19	0.035	0.007	0.44	0.157	0.007	0.69	0.396	0.006	0.94	0.742	0.003
0.2	0.032	0.007	0.45	0.169	0.007	0.7	0.417	0.006	0.95	0.758	0.003
0.21	0.036	0.007	0.46	0.170	0.007	0.71	0.427	0.006	0.96	0.776	0.003
0.22	0.049	0.007	0.47	0.188	0.007	0.72	0.439	0.006	0.97	0.789	0.003
0.23	0.031	0.007	0.48	0.202	0.007	0.73	0.446	0.006	0.98	0.808	0.002
0.24	0.056	0.007	0.49	0.216	0.007	0.74	0.456	0.006	0.99	0.826	0.002

Table 2: Data for Figure 3.

θ	Sol.	δ
0.000	-0.012	0.007
0.063	0.047	0.007
0.126	0.062	0.007
0.188	0.099	0.007
0.251	0.130	0.007
0.314	0.150	0.007
0.377	0.174	0.007
0.440	0.187	0.007
0.503	0.206	0.007
0.565	0.225	0.007
0.628	0.242	0.007
0.691	0.226	0.007
0.754	0.243	0.007
0.817	0.258	0.007
0.880	0.230	0.007
0.942	0.241	0.007
1.005	0.237	0.007
1.068	0.210	0.007
1.131	0.178	0.007
1.194	0.169	0.007
1.257	0.154	0.007
1.319	0.121	0.007
1.382	0.097	0.007
1.445	0.063	0.007
1.508	0.035	0.007

θ	Sol.	δ
1.571	-0.009	0.007
1.634	-0.031	0.007
1.696	-0.049	0.007
1.759	-0.086	0.007
1.822	-0.126	0.007
1.885	-0.149	0.007
1.948	-0.175	0.007
2.011	-0.193	0.007
2.073	-0.223	0.007
2.136	-0.230	0.007
2.199	-0.236	0.007
2.262	-0.240	0.007
2.325	-0.254	0.007
2.388	-0.248	0.007
2.450	-0.243	0.007
2.513	-0.243	0.007
2.576	-0.231	0.007
2.639	-0.211	0.007
2.702	-0.186	0.007
2.765	-0.177	0.007
2.827	-0.148	0.007
2.890	-0.112	0.007
2.953	-0.080	0.007
3.016	-0.067	0.007
3.079	-0.028	0.007

θ	Sol.	δ
3.142	-0.010	0.007
3.204	0.030	0.007
3.267	0.064	0.007
3.330	0.092	0.007
3.393	0.110	0.007
3.456	0.143	0.007
3.519	0.164	0.007
3.581	0.193	0.007
3.644	0.212	0.007
3.707	0.223	0.007
3.770	0.238	0.007
3.833	0.258	0.007
3.896	0.248	0.007
3.958	0.243	0.007
4.021	0.244	0.007
4.084	0.251	0.007
4.147	0.220	0.007
4.210	0.214	0.007
4.273	0.183	0.007
4.335	0.188	0.007
4.398	0.142	0.007
4.461	0.116	0.007
4.524	0.100	0.007
4.587	0.059	0.007
4.650	0.032	0.007

θ	Sol.	δ
4.712	-0.003	0.007
4.775	-0.024	0.007
4.838	-0.057	0.007
4.901	-0.084	0.007
4.964	-0.128	0.007
5.027	-0.150	0.007
5.089	-0.169	0.007
5.152	-0.182	0.007
5.215	-0.214	0.007
5.278	-0.228	0.007
5.341	-0.227	0.007
5.404	-0.251	0.007
5.466	-0.244	0.007
5.529	-0.260	0.007
5.592	-0.257	0.007
5.655	-0.233	0.007
5.718	-0.222	0.007
5.781	-0.202	0.007
5.843	-0.179	0.007
5.906	-0.184	0.007
5.969	-0.143	0.007
6.032	-0.128	0.007
6.095	-0.096	0.007
6.158	-0.058	0.007
6.220	-0.037	0.007

Table 3: Data for Figure 4. $E[N]$ denotes expected number of steps until exit.

r	$E[N]$	δ
0	5211	117
0.01	5093	119
0.02	4952	104
0.03	4907	107
0.04	5106	113
0.05	5007	114
0.06	4984	113
0.07	5032	108
0.08	5270	120
0.09	5029	118
0.1	4818	103
0.11	5052	113
0.12	5071	118
0.13	5171	122
0.14	4856	115
0.15	5143	119
0.16	4984	115
0.17	5073	120
0.18	4866	111
0.19	4769	111
0.2	4984	114
0.21	4610	110
0.22	4965	116
0.23	4601	112
0.24	4866	123

r	$E[N]$	δ
0.25	4735	110
0.26	4767	110
0.27	4524	109
0.28	4644	116
0.29	4617	112
0.3	4570	114
0.31	4430	102
0.32	4188	100
0.33	4506	113
0.34	4188	102
0.35	4487	113
0.36	4307	109
0.37	4230	109
0.38	4460	113
0.39	4470	118
0.4	4129	105
0.41	4049	104
0.42	4155	110
0.43	4023	111
0.44	4098	114
0.45	4099	110
0.46	4060	107
0.47	4017	114
0.48	3705	102
0.49	3641	105

r	$E[N]$	δ
0.5	3913	111
0.51	3638	102
0.52	3574	100
0.53	3490	105
0.54	3617	115
0.55	3585	111
0.56	3456	105
0.57	3460	108
0.58	3585	105
0.59	3233	109
0.6	3375	105
0.61	2997	99
0.62	3095	106
0.63	3143	107
0.64	2961	101
0.65	3135	108
0.66	2785	98
0.67	2711	95
0.68	2667	103
0.69	2630	98
0.7	2559	96
0.71	2664	105
0.72	2453	101
0.73	2294	87
0.74	2275	98

r	$E[N]$	δ
0.75	2251	101
0.76	2053	92
0.77	2088	90
0.78	2027	86
0.79	1892	86
0.8	1771	84
0.81	1847	90
0.82	1720	90
0.83	1683	83
0.84	1480	74
0.85	1467	82
0.86	1297	75
0.87	1242	77
0.88	1116	66
0.89	1020	64
0.9	993	66
0.91	835	57
0.92	753	57
0.93	662	53
0.94	648	58
0.95	512	46
0.96	363	38
0.97	281	35
0.98	220	31
0.99	119	22

Table 4: Data for Figure 5. $E[N]$ denotes expected number of steps until exit.

Δt	$E[N]$	δ
1.00E-01	5.9	0.4
3.16E-02	18.3	1.5
1.00E-02	38.3	3.3
3.16E-03	118	10
1.00E-03	340	29
3.16E-04	1104	95
1.00E-04	3672	370
3.16E-05	11177	909
1.00E-05	36654	3072
3.16E-06	146367	14641
1.00E-06	375674	33530

Table 5: Data for Figure 6.

Δt	Sol.	δ
1.00E-01	0.256	0.010
3.16E-02	0.238	0.008
1.00E-02	0.242	0.008
3.16E-03	0.237	0.007
1.00E-03	0.245	0.007
3.16E-04	0.249	0.007
1.00E-04	0.257	0.007
3.16E-05	0.242	0.007
1.00E-05	0.245	0.007

Table 6: Data for Figure 8.

θ	Sol.	δ	θ	Sol.	δ	θ	Sol.	δ	θ	Sol.	δ
0.00	1.581	0.006	1.57	1.401	0.003	3.14	1.572	0.006	4.71	1.404	0.003
0.06	1.596	0.006	1.63	1.386	0.003	3.20	1.596	0.006	4.78	1.387	0.003
0.13	1.614	0.006	1.70	1.374	0.004	3.27	1.620	0.006	4.84	1.372	0.004
0.19	1.643	0.006	1.76	1.357	0.004	3.33	1.648	0.006	4.90	1.361	0.004
0.25	1.656	0.006	1.82	1.349	0.004	3.39	1.660	0.006	4.96	1.354	0.004
0.31	1.659	0.006	1.88	1.341	0.005	3.46	1.665	0.006	5.03	1.344	0.005
0.38	1.683	0.005	1.95	1.336	0.005	3.52	1.680	0.005	5.09	1.341	0.005
0.44	1.686	0.005	2.01	1.337	0.005	3.58	1.680	0.005	5.15	1.323	0.005
0.50	1.689	0.005	2.07	1.333	0.005	3.64	1.676	0.005	5.22	1.336	0.005
0.57	1.682	0.004	2.14	1.325	0.005	3.71	1.675	0.004	5.28	1.331	0.006
0.63	1.670	0.004	2.20	1.334	0.006	3.77	1.685	0.004	5.34	1.333	0.006
0.69	1.671	0.003	2.26	1.332	0.006	3.83	1.670	0.003	5.40	1.338	0.006
0.75	1.663	0.003	2.32	1.338	0.006	3.90	1.660	0.003	5.47	1.354	0.006
0.82	1.648	0.003	2.39	1.340	0.006	3.96	1.649	0.003	5.53	1.349	0.006
0.88	1.633	0.002	2.45	1.348	0.006	4.02	1.634	0.002	5.59	1.359	0.006
0.94	1.615	0.002	2.51	1.377	0.006	4.08	1.614	0.002	5.65	1.373	0.006
1.01	1.593	0.002	2.58	1.390	0.006	4.15	1.596	0.002	5.72	1.384	0.006
1.07	1.575	0.001	2.64	1.400	0.006	4.21	1.573	0.001	5.78	1.405	0.006
1.13	1.551	0.001	2.70	1.412	0.006	4.27	1.554	0.001	5.84	1.427	0.006
1.19	1.531	0.001	2.76	1.450	0.007	4.34	1.533	0.001	5.91	1.427	0.006
1.26	1.507	0.001	2.83	1.454	0.007	4.40	1.507	0.001	5.97	1.465	0.007
1.32	1.486	0.001	2.89	1.478	0.007	4.46	1.487	0.002	6.03	1.480	0.007
1.38	1.460	0.002	2.95	1.514	0.007	4.52	1.465	0.002	6.09	1.505	0.007
1.45	1.441	0.002	3.02	1.532	0.007	4.59	1.442	0.002	6.16	1.529	0.007
1.51	1.423	0.003	3.08	1.562	0.007	4.65	1.425	0.003	6.22	1.559	0.007

Table 7: Data for Figure 9.

r	Sol.	δ	r	Sol.	δ	r	Sol.	δ	r	Sol.	δ
0	1.546	0.001	0.25	1.505	0.004	0.5	1.347	0.006	0.75	0.975	0.004
0.01	1.546	0.001	0.26	1.498	0.004	0.51	1.334	0.006	0.76	0.953	0.004
0.02	1.543	0.001	0.27	1.493	0.004	0.52	1.326	0.006	0.77	0.944	0.004
0.03	1.542	0.001	0.28	1.483	0.004	0.53	1.304	0.006	0.78	0.927	0.004
0.04	1.543	0.001	0.29	1.494	0.004	0.54	1.299	0.006	0.79	0.912	0.004
0.05	1.542	0.001	0.3	1.471	0.004	0.55	1.289	0.006	0.8	0.905	0.004
0.06	1.543	0.001	0.31	1.476	0.004	0.56	1.276	0.006	0.81	0.884	0.003
0.07	1.540	0.002	0.32	1.468	0.005	0.57	1.256	0.006	0.82	0.866	0.003
0.08	1.541	0.002	0.33	1.464	0.005	0.58	1.223	0.006	0.83	0.847	0.003
0.09	1.539	0.002	0.34	1.463	0.005	0.59	1.229	0.006	0.84	0.829	0.003
0.1	1.539	0.002	0.35	1.447	0.005	0.6	1.214	0.006	0.85	0.815	0.003
0.11	1.536	0.002	0.36	1.441	0.005	0.61	1.203	0.006	0.86	0.805	0.003
0.12	1.536	0.002	0.37	1.442	0.005	0.62	1.180	0.006	0.87	0.793	0.003
0.13	1.532	0.002	0.38	1.454	0.005	0.63	1.163	0.006	0.88	0.777	0.002
0.14	1.535	0.002	0.39	1.433	0.005	0.64	1.146	0.005	0.89	0.762	0.002
0.15	1.529	0.002	0.4	1.429	0.005	0.65	1.132	0.005	0.9	0.749	0.002
0.16	1.536	0.003	0.41	1.422	0.006	0.66	1.114	0.005	0.91	0.736	0.002
0.17	1.521	0.003	0.42	1.416	0.006	0.67	1.105	0.005	0.92	0.727	0.002
0.18	1.515	0.003	0.43	1.404	0.006	0.68	1.077	0.005	0.93	0.710	0.002
0.19	1.518	0.003	0.44	1.405	0.006	0.69	1.072	0.005	0.94	0.695	0.002
0.2	1.514	0.003	0.45	1.389	0.006	0.7	1.050	0.005	0.95	0.683	0.002
0.21	1.516	0.003	0.46	1.390	0.006	0.71	1.037	0.005	0.96	0.670	0.001
0.22	1.511	0.003	0.47	1.374	0.006	0.72	1.015	0.004	0.97	0.662	0.001
0.23	1.511	0.003	0.48	1.371	0.006	0.73	1.007	0.004	0.98	0.649	0.001
0.24	1.509	0.004	0.49	1.349	0.006	0.74	0.989	0.004	0.99	0.637	0.001

Table 8: Data for Figure 10.

r	Sol.	δ	r	Sol.	δ	r	Sol.	δ	r	Sol.	δ
0.000	1.542	0.001	0.354	1.602	0.002	0.707	1.758	0.004	1.061	2.011	0.004
0.014	1.546	0.001	0.368	1.600	0.002	0.721	1.771	0.004	1.075	2.019	0.004
0.028	1.545	0.001	0.382	1.606	0.002	0.735	1.777	0.004	1.089	2.033	0.004
0.042	1.546	0.001	0.396	1.613	0.002	0.750	1.783	0.004	1.103	2.045	0.004
0.057	1.545	0.001	0.410	1.620	0.002	0.764	1.794	0.004	1.117	2.050	0.004
0.071	1.546	0.001	0.424	1.619	0.003	0.778	1.799	0.004	1.131	2.063	0.004
0.085	1.548	0.001	0.438	1.630	0.003	0.792	1.809	0.004	1.146	2.075	0.004
0.099	1.549	0.001	0.453	1.633	0.003	0.806	1.819	0.004	1.160	2.088	0.004
0.113	1.549	0.001	0.467	1.634	0.003	0.820	1.831	0.004	1.174	2.103	0.004
0.127	1.553	0.001	0.481	1.647	0.003	0.834	1.839	0.004	1.188	2.116	0.003
0.141	1.552	0.001	0.495	1.653	0.003	0.849	1.859	0.004	1.202	2.129	0.003
0.156	1.556	0.001	0.509	1.657	0.003	0.863	1.860	0.004	1.216	2.141	0.003
0.170	1.556	0.001	0.523	1.665	0.003	0.877	1.873	0.004	1.230	2.155	0.003
0.184	1.560	0.002	0.537	1.672	0.003	0.891	1.874	0.004	1.245	2.163	0.003
0.198	1.560	0.002	0.552	1.682	0.003	0.905	1.889	0.004	1.259	2.173	0.003
0.212	1.563	0.002	0.566	1.680	0.003	0.919	1.898	0.004	1.273	2.187	0.003
0.226	1.566	0.002	0.580	1.691	0.003	0.933	1.905	0.004	1.287	2.207	0.003
0.240	1.567	0.002	0.594	1.696	0.003	0.948	1.923	0.004	1.301	2.218	0.003
0.255	1.571	0.002	0.608	1.707	0.003	0.962	1.931	0.004	1.315	2.232	0.003
0.269	1.575	0.002	0.622	1.712	0.003	0.976	1.943	0.004	1.329	2.243	0.002
0.283	1.580	0.002	0.636	1.725	0.003	0.990	1.952	0.004	1.344	2.255	0.002
0.297	1.584	0.002	0.651	1.728	0.003	1.004	1.965	0.004	1.358	2.269	0.002
0.311	1.587	0.002	0.665	1.735	0.003	1.018	1.976	0.004	1.372	2.287	0.002
0.325	1.595	0.002	0.679	1.745	0.003	1.032	1.978	0.004	1.386	2.298	0.002
0.339	1.594	0.002	0.693	1.747	0.004	1.047	1.995	0.004	1.400	2.318	0.001

Table 9: Data for Figure 11.

Δt	θ	Sol.	δ
1.00E-04	0.00	1.589	0.021
1.00E-04	0.31	1.693	0.018
1.00E-04	0.63	1.706	0.013
1.00E-04	0.94	1.642	0.006
1.00E-04	1.26	1.530	0.004
1.00E-04	1.57	1.434	0.010
1.00E-04	1.88	1.367	0.015
1.00E-04	2.20	1.347	0.018
1.00E-04	2.51	1.403	0.021
1.00E-04	2.83	1.480	0.021
1.00E-04	3.14	1.603	0.021
1.00E-04	3.46	1.697	0.018
1.00E-04	3.77	1.665	0.013
1.00E-04	4.08	1.647	0.006
1.00E-04	4.40	1.531	0.004
1.00E-04	4.71	1.436	0.010
1.00E-04	5.03	1.361	0.015
1.00E-04	5.34	1.355	0.019
1.00E-04	5.65	1.357	0.020
1.00E-04	5.97	1.486	0.021

Δt	θ	Sol.	δ
1.00E-06	0.00	1.578	0.020
1.00E-06	0.31	1.654	0.017
1.00E-06	0.63	1.687	0.012
1.00E-06	0.94	1.609	0.006
1.00E-06	1.26	1.502	0.004
1.00E-06	1.57	1.410	0.009
1.00E-06	1.88	1.349	0.014
1.00E-06	2.20	1.319	0.018
1.00E-06	2.51	1.353	0.019
1.00E-06	2.83	1.429	0.020
1.00E-06	3.14	1.581	0.021
1.00E-06	3.46	1.652	0.018
1.00E-06	3.77	1.677	0.012
1.00E-06	4.08	1.611	0.006
1.00E-06	4.40	1.507	0.004
1.00E-06	4.71	1.394	0.009
1.00E-06	5.03	1.326	0.014
1.00E-06	5.34	1.324	0.018
1.00E-06	5.65	1.373	0.020
1.00E-06	5.97	1.446	0.020