THESIS

NUMERICAL SOLUTION OF THE BLACK-SCHOLES EQUATION USING FINITE ELEMENT METHODS

Submitted by

Tyler Anderson

Department of Mathematics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2023

Master's Committee:

Advisor: Wolfgang Bangerth

David Aristoff
Tianyang Wang

ABSTRACT

NUMERICAL SOLUTION OF THE BLACK-SCHOLES EQUATION USING FINITE ELEMENT METHODS

The Black-Scholes model is a well known model for pricing financial options. This model takes the form of a partial differential equation (PDE) that, surprisingly, is deterministic. In the special case where the option only has one single underlying asset, what is called the one dimensional version of the Black-Scholes model, there exists an analytical solution. In higher dimensions, however, there is no such analytical solution. This higher dimensional version refers to what is called a Basket-Case Option. This means that to get a solution to this Basket-Case Option PDE, one must employ numerical methods. This thesis will first discuss the stochastic calculus theory necessary to derive the Black-Scholes model, then will explain in detail the time and space discretization used to solve the PDE using a Finite Element Method (FEM). Finally, this thesis will explain some of the results and convergence of this numerical solution.

ACKNOWLEDGEMENTS

I would like to thank my Advisor, Dr. Wolfgang Bangerth, for his support, encouragement and guidance. I would like to thank him especially for letting me explore while learning. I would also like to thank my wife, Montse, for her unending support and belief in me during this journey. I could not have done this without her faith in me and for her sacrifices. Finally, I would like to thank my family for their support. I would like to thank them for believing in me.

DEDICATION

I would like to dedicate this to my Granny and Papa. This '100' is for you.

TABLE OF CONTENTS

ABSTRACT ii
ACKNOWLEDGEMENTS iii
DEDICATION iv
LIST OF TABLES vii
LIST OF FIGURES viii

Chapter 1 Introduction 1

Chapter 2 Stochastic Background 5
2.1 The Wiener Process 5
2.2 Stochastic Integral 6
2.2.1 Basic Stochastic Lemmas 8
2.2.2 Quadratic Variation 11
2.2.3 Integrate Wiener Process 16
2.2.4 General Stochastic Integral 20
2.3 Stochastic Differential Equations 23
2.3.1 Basic Stochastic Differentials 25
2.3.2 Basic Multi-Dimensional Stochastic Differentials 29
2.3.3 Product Rule 35
2.3.4 Itô's Lemma 41
2.4 Chapter Summary 47

Chapter 3 Black-Scholes Model 48
3.1 One-Dimensional Black-Scholes Model 49
3.1.1 Delta Hedging 49
3.1.2 One-Dimensional Derivation 50
3.1.3 One-Dimensional Boundary and Initial Conditions 52
3.1.4 One-Dimensional Model Statement 54
3.2 Basket-Case Black-Scholes Model 54
3.2.1 Asset Modification 55
3.2.2 Multi-Asset Vector Model 61
3.2.3 Basket-Case Derivation 63
3.2.4 Left-Side Basket-Case Boundary Conditions 66
3.2.5 Basket-Case Boundary and Initial Conditions 68
3.2.6 Three-Dimensional Basket-Case Black-Scholes Model 69
3.3 Chapter Summary 71

Chapter 4 Numerical Solution 72
4.1 Truncation of the Spatial Domain 72
4.2 Conservative Formulation 74
4.3 Weak Formulation 78

4.3.1 Time Reversal . . . 78
4.3.2 Time Domain Discretization . . . 79
4.3.3 Spatial Discretization . . . 81
4.4 Complete Model . . . 87
4.5 Peclet Number . . . 94
4.6 Chapter Summary . . . 95

Chapter 5 Results . . . 98
5.1 Three-Dimensional Solution . . . 98
5.1.1 Iso-Surfaces . . . 99
5.1.2 Two-Dimensional Slice . . . 102
5.2 Method of Manufactured Solutions . . . 102
5.2.1 Convergence and Error . . . 104

Chapter 6 Summary and Conclusion . . . 107

References . . . 109

LIST OF TABLES

5.1 Error Table. *Table showing relationship between the refinement of the mesh and the global errors with respect to various norms and semi-norms.* 105
5.2 Convergence Rate Table. *Table showing the convergence rates with respect to the H^1 semi-norm and the L^2 norm.* 106

LIST OF FIGURES

2.1 Simulated Realization of Wiener Process 7

4.1 Modified Peclet Number 2-D. 96

5.1 Iso-Surfaces of solution to 3D Basket-Case Model. *Iso-surfaces of the solution to a three dimensional Basket-Case option with strike price* $K = \$0.5$. 99
5.2 Two-Dimensional Slice of Domain. 101
5.3 Elevated Slice at Initial Time and Expiration Time. *Two dimensional elevated slices with* $K = \$0.5$. 102
5.4 MMS Solution Iso-Surfaces. *Iso-surfaces for* $f(S_1, S_2, S_3, \tau) = 6 - S_1^2 - S_2^2 - S_3^2 - \tau^2$. 103

Chapter 1

Introduction

The Black-Scholes Equation is a famous equation used to model the value of financial options [8]. There are two types of financial options that are generally talked about, those being the *Call Option* and the *Put Option* [1]. In this paper, I will discuss a slight variant of the call option, called a *European Call Option*.

A **European Call Option** is a contract in essence. It is a contract between a buyer and seller of an asset where at a set time in the future, or the *expiration date*, the buyer has the *right* to buy the asset, but not the obligation, at an agreed upon price, called the *strike price*[4]. This may seem like a fairly abstract financial concept, but it is essentially like purchasing insurance on a financial asset.

An example of this in the real world is in the airline industry. Airline companies will frequently purchase an option for purchasing jet fuel. The reason for this is that fuel prices change rapidly, but airline companies book reservations for flights well into the future. This means that they have to make a good estimate for how much it will cost to fly an aircraft on a specific day in the future so that they can still make money off of the tickets that they are selling. One way that this is done is buy purchasing an option for jet fuel. This allows for them to have an upper limit on the fuel prices at a specific day. If the fuel prices go above the strike price, then they will exercise the option and purchase the fuel at the strike price. If the fuel prices fall below the strike price, then they will not exercise the option, and will simply purchase the fuel at the market price for that day. This allows the airline companies to create plane ticket prices, because they have a reasonable estimate for the price of the fuel on a specific day.

This way of protecting, or insuring, yourself against market fluctuations is called *hedging* [4].

Market fluctuations are what has prompted research into trying to accurately predict prices of financial assets. Because call options derive their value from the value of their underlying asset, such as a collection of stocks, then it is of particular interest to accurately assign a value to these

options. By assigning a value to the option, I mean assigning an appropriate premium price. If these prices are too high, then this would cause the option to be overvalued and unfairly favor the seller of the option. If these prices are too low, then this would cause the option to be undervalued and unfairly favor the holder, or buyer, of the option.

Fischer Black and Myron Scholes introduced a model that would attach a value to these call options [8]. The model that they introduced was a partial differential equation (PDE) with a unique solution that assigns a value to the option [8].

Now, although a financial call option can describe a contract to buy any kind of asset, in this thesis I will only consider a stock call option. This is simply a call option where the underlying asset is a stock.

The Black-Scholes Model PDE is a linear advection-diffusion-reaction equation. For the one dimensional case, or the case where our portfolio only contains one stock, this PDE actually has an analytical solution. This solution is realized by doing several changes of variables and arriving at the one-dimensional heat equation, for which there is a known analytical solution [8]. Although there is an analytical solution to the one-dimensional problem, it is still a problem of interest to look at solving numerically. This is because the PDE is advective-diffusive-reactive, so one needs to handle the equation carefully to produce a stable and accurate solution, and also because the one-dimensional solution sheds some light on how to solve the multi-dimensional problem. By multi-dimensional, I am referring to what is called a **Basket-Case Option**, or an option with more than one underlying stock. Diffusive equations are generally forgiving when it comes to the time and space discretization, but because of the advective nature of this problem one needs to be careful.

Numerical solutions to PDEs are useful when it comes to solving problems that may or may not have analytical solutions. They are useful, of course, when there are no analytical solutions available because we are still able to produce a solution, or at least an approximation of one.

This PDE has previously been solved numerically through the use of Finite Difference Methods [4], and through the use of Monte-Carlo Simulations [19]. Although this PDE has also been solved

using the Finite Element Method [10], I intend to develop my own understanding of the Finite Element solution in this thesis. This thesis then, will attempt to solve this PDE with the Finite Element Method.

In this thesis, I will first discuss the underlying stochastic calculus that is necessary to derive the Black-Scholes model for the Basket-Case Option. Then, with the problem well defined, I will explain the time and space discretization I will use in my Finite Element Method solution. Finally, I will discuss the convergence and accuracy of this numerical solution using the Method of Manufactured Solutions (MMS) [24].

In Chapter 2, I will give an overview of the stochastic calculus necessary to derive the Black-Scholes Basket-Case Option PDE. This will include a discussion on the basic building block of Stochastic Calculus: the Wiener Process [26]. Then, with this process understood, I can describe some important lemmas and theorems in Stochastic Calculus that will be important to deriving the Black-Scholes Basket-Case Option PDE, most notably Itô's Lemma [26].

In Chapter 3, I will use the basics of Stochastic Calculus discussed in Chapter 2 to derive the Black-Scholes Model in one dimension and in multi-dimensions. The latter case of course being called the Basket-Case Option Model.

In Chapter 4, I will describe the discretization I will use for solving the Basket-Case Option PDE. I will discuss how to handle the diffusive and advective terms separately using an Implicit-Explicit (IMEX) method that will maintain a high order of accuracy in time and space [11]. Implicit schemes are generally more accurate, but suffer the consequence of being more computationally expensive and require certain properties of the problem [7]. One property that implicit schemes require is the cheap computation of a matrix inverse. Because diffusion operators generally result in cheaply invertable matrices, these types of operators are well suited for implicit schemes. Explicit schemes, on the other hand, require much less regularity in the problem and can be quite cheap to compute, but they are conditionally stable [7]. Since advective operators typically don't result in matrices that have cheaply computable inverses, then explicit schemes are necessary for solving problems with this operator. It does mean, however, that care must be taken when it comes to the

stability of the solution when solving with an explicit scheme. In this chapter then, I will describe the IMEX method that I will be using in my discretization. This will involve using an implicit scheme on the diffusive operator and an explicit scheme on the advective operator.

Finally, in Chapter 5, I will discuss the results of the program I wrote to compute the solution to the Black-Scholes Basket-Case Option Model. This will include diagrams of the solution and discussions on what these diagrams represent, as well as a discussion on the convergence of my solution with the discretization that was described in Chapter 4.

My thesis will conclude with a summary of these chapters and my results.

Chapter 2

Stochastic Background

The Black-Scholes Model is about pricing financial options. Recall that a call option is a contract to potentially buy some amount of stocks at a date in the future. Thus, in order to understand how to price a call option, I must first consider how the price of the underlying stock behaves.

When trying to determine how to model the price of a particular stock, there are so many factors that it becomes necessary to model the general behavior and then add in some sort of randomness. This randomness in the price of the stock means that it will be necessary to have a basic understanding of random processes that evolve in time, or **Stochastic Processes**.

This chapter will focus on the stochastic knowledge necessary to understand how the price of the underlying stocks behave, and subsequently, the value of the call option. The ultimate goal of this chapter is to understand an important stochastic calculus theorem, called Itô's Lemma [17].

Understanding Itô's lemma will require some knowledge about the subject of Stochastic Calculus. This is, as the name implies, calculus on stochastic processes.

I will start out by discussing the most basic element I will need in order to understand stochastic processes, the Wiener Process. This concept will allow me to construct the stochastic integral through concepts similar to Riemann Sums. Then, I will define the stochastic differential and talk about some basic Stochastic Calculus theorems. This, will finally lead me to the most useful theorem in my thesis, which is Itô's Lemma. This lemma will be crucial in deriving the Black-Scholes PDE.

2.1 The Wiener Process

Stock prices are extremely difficult to predict and model [14]. Because of this, in my thesis I will be using a model with some amount of randomness to describe the value of these stocks. As I stated in the introduction above, this means that I first need to understand one of the more simple, but important, stochastic processes: the **Wiener Process**. This stochastic process is important

because it will allow me to talk about more complicated stochastic processes, such as the price of a stock.

I will begin by giving the definition of a Wiener Process.

Definition 1. A real-valued stochastic process $W(\cdot)$ is said to be a *Wiener Process* if

1. $W(0) = 0 \quad$ *a.s.*,
2. $W(t) - W(s) \sim N(0, t-s) \quad \forall t \geq s \geq 0$,
3. for all $0 \leq t_1 \leq t_2 \leq ... \leq t_n$, the random variables $W(t_1), W(t_2) - W(t_1), ..., W(t_n) - W(t_{n-1})$ are independent (independent increments).

In this definition, *a.s.* means *almost surely* and $N(m, s)$ means *normally distributed* with mean m and standard deviation s. These definitions can be explored further in [15].

The first part of the definition simply means that there are only a finite number of times that the random event $W(0)$ will not be equal to 0. The second part of the definition describes the distribution of the difference of a change in the random variable in some unit of time. The last part of the definition simply describes that these differences are independent, so long as the time intervals are disjoint.

In Figure 2.1, I have simulated a Wiener Process to show how these stochastic processes evolve. In this figure, I show 5 realizations of this stochastic process.

This is an important stochastic process, although simple, because it is going to allow me to discuss the next important topic: The Stochastic Integral.

2.2 Stochastic Integral

The Black-Scholes Model is a partial differential equation that describes the value of a call option. This option, since it depends upon the value of its underlying stocks, will also have its value be described by a stochastic process. It will be important, then, to be able to talk about differential equations involving stochastic processes. Unfortunately, the Wiener Process is *nowhere-*

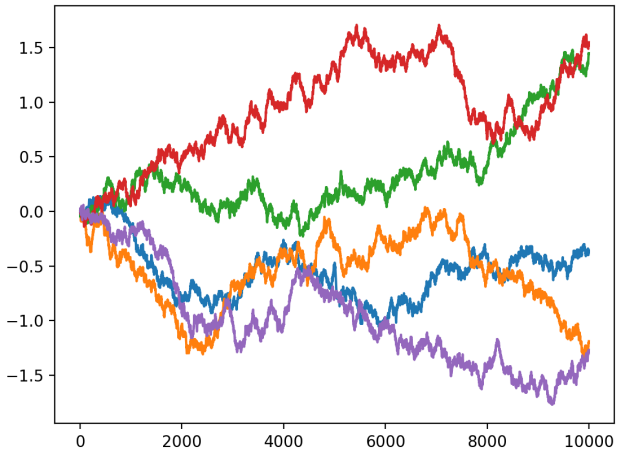


Figure 2.1: Simulated Realization of Wiener Process

differentiable [15]. This will obviously be a problem if I hope to talk about differential equations involving stochastic processes, often called Stochastic Differential Equations (SDEs).

What I will do then, is to define the stochastic integral first. Surprisingly, this is somewhat easier to define. This is somewhat backwards to how "*regular*" calculus is taught, but because of the nowhere-differentiability of the Wiener Process, this is the path that I will take. Then, I will use the definition of the stochastic integral to define the *stochastic differential*.

For this thesis, I will limit the discussion of a stochastic integral to integrating with respect to a Wiener Process. Mathematically, I am trying to understand the integral

$$\int_0^T G dW, \tag{2.1}$$

where G is a stochastic process. In fact, G is not just any ordinary stochastic process, but is a "well behaved" stochastic process. I am putting quotes around this because there are several restrictions on G that must be true that I will not be going into in this thesis. For more information, see [15]. Loosely, the restrictions are:

1. For each time $t \geq 0$, $G(t)$ must only depend upon the information available up to time t from a Wiener Process and
2. $E\left[\int_0^T G^2 dt\right] < \infty$.

In order to understand the integral in Equation (2.1), it seems reasonable to consider trying a Riemann-like sum similar to how integrals were constructed in "ordinary" calculus.

I will start this section by giving some basic stochastic lemmas, and then defining the stochastic integral in Equation (2.1).

2.2.1 Basic Stochastic Lemmas

This section will describe some basic stochastic lemmas that will be necessary in defining the stochastic integral in Equation (2.1). In particular, these will be useful in showing that a very important property of Wiener Processes holds: *Quadratic Variation*. The first of these lemmas talks about the distribution of a normalized difference of Wiener Processes. The second lemma talks about the expectation of a random variable that is related to the normalized difference random variable in the first lemma.

Lemma 1. *Let* $[a, b]$ *be an interval in* $[0, \infty)$ *and suppose that this interval has been partitioned into* m_n *equally sized intervals such that*

$$P^n := \{a = t_0^n < t_1^n < ... < t_{m_n}^n = b\}$$

are partitions of $[a, b]$. *Also suppose that* $W(t)$ *is a Wiener Process and that*

$$\tau_k := \lambda t_{k+1}^n + (1 - \lambda)\, t_k^n,$$

where $\lambda \in (0, 1)$. *Finally let*

$$Y_k^n := \frac{W(\tau_k) - W(t_k^n)}{\sqrt{\lambda \left(t_{k+1}^n - t_k^n\right)}}.$$

Then $Y_k^n \sim N(0, 1)$.

Proof. First, I will define a new random variable:

$$R_k^n := W(\tau_k) - W(t_k^n) \sim N(0, \tau_k - t_k^n).$$

This is the numerator of the random variable Y_k^n. Now, because $\tau_k = \lambda t_{k+1}^n + (1-\lambda)t_k^n \geq t_k^n$, I have

$$R_k^n \sim N\left(0, \lambda\left(t_{k+1}^n - t_k^n\right)\right)$$

by Definition 1. The variance came from the simplification:

$$\begin{aligned}\tau_k - t_k^n &= \lambda t_{k+1}^n + (1-\lambda)t_k^n - t_k^n \\ &= \lambda\left(t_{k+1}^n - t_k^n\right).\end{aligned}$$

This was all to make the next few calculations a bit simpler to do. Notice that:

$$Y_k^n = \frac{R_k^n}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)}}$$

because of my definition of R_k^n. Now, the denominator of Y_k^n is simply a constant, and not a random process. This means then, that the distribution of Y_k^n must be a normal distribution since R_k^n is normally distributed. So, all that is left to show is the mean and variance of Y_k^n. First, for the mean I have:

$$\begin{aligned}E\left[Y_k\right] &= E\left[\frac{R_k^n}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)}}\right] \\ &= \frac{E\left[R_k^n\right]}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)}} \\ &= \frac{0}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)}} \\ &= 0.\end{aligned}$$

Thus, I have that the mean of Y_k^n is 0. All that is left then, is to show that the variance is 1. First, recall that

$$\text{var}\left(Y_k^n\right) = E\left[\left(Y_k^n\right)^2\right] - \left(E\left[Y_k^n\right]\right)^2.$$

This is sometimes taken as the definition of variance, but I will leave the reader to read [13] for more information.

Now, using the definition of variance from above, I have:

$$\begin{aligned}
\operatorname{var}\left(Y_k^n\right) &= E\left[\left(Y_k^n\right)^2\right] - \left(E\left[Y_k^n\right]\right)^2 \\
&= E\left[\left(\frac{R_k^n}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)^2}}\right)^2\right] - \left(E\left[\frac{R_k^n}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)^2}}\right]\right)^2 \\
&= E\left[\frac{\left(R_k^n\right)^2}{\lambda\left(t_{k+1}^n - t_k^n\right)}\right] - \left(\frac{E\left[R_k^n\right]}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)^2}}\right)^2 \\
&= \frac{E\left[\left(R_k^n\right)^2\right] - \left(E\left[R_k^n\right]\right)^2}{\lambda\left(t_{k+1}^n - t_k^n\right)} \\
&= \frac{\operatorname{var}\left(R_k^n\right)}{\lambda\left(t_{k+1}^n - t_k^n\right)}.
\end{aligned}$$

I showed at the beginning of this proof that $\operatorname{var}\left(R_k^n\right) = \lambda\left(t_{k+1}^n - t_k^n\right)$. Thus, I can go ahead and substitute this into the numerator of my expression to see that

$$\begin{aligned}
\operatorname{var}\left(Y_k^n\right) &= \frac{\operatorname{var}\left(R_k^n\right)}{\lambda\left(t_{k+1}^n - t_k^n\right)} \\
&= \frac{\lambda\left(t_{k+1}^n - t_k^n\right)}{\lambda\left(t_{k+1}^n - t_k^n\right)} \\
&= 1.
\end{aligned}$$

Thus, I have shown that $Y_k^n \sim N(0,1)$. □

Now that I have shown the distribution for a normalized difference of Wiener Processes, I can move on to the second important lemma in this section.

Lemma 2. *Suppose that the interval $[a,b] \in [0,\infty)$ is defined the same as in Lemma 1 and let Y_k^n be defined similarly as well. Then,*

$$E\left[\left(Y_k^n - 1\right)^2\right] = 2.$$

Proof. I will start by recalling that the variance of $Y_k^n - 1$ is given by:

$$\operatorname{var}\left(Y_k^n - 1\right) = E\left[\left(Y_k^n - 1\right)^2\right] - \left(E\left[Y_k^n - 1\right]\right)^2.$$

Rearranging these terms, I see that

$$E\left[\left(Y_k^n - 1\right)^2\right] = \operatorname{var}\left(Y_k^n - 1\right) + \left(E\left[Y_k^n - 1\right]\right)^2.$$

The first part of this sum will simply be equal to $\operatorname{var}\left(Y_k^n\right)$ because 1 is a constant number [25]. Thus, I have

$$E\left[\left(Y_k^n - 1\right)^2\right] = \operatorname{var}\left(Y_k^n\right) + \left(E\left[Y_k^n\right] - 1\right)^2.$$

This is because the expectation operator is linear. [25]. Expanding the second term, I get

$$\begin{aligned} E\left[\left(Y_k^n - 1\right)^2\right] &= \operatorname{var}\left(Y_k^n\right) + \left(E\left[Y_n^k\right]\right)^2 - 2E\left[Y_n^k\right] + 1 \\ &= 1 + 1 \\ &= 2. \end{aligned}$$

This follows from Lemma 1 that $Y_k^n \sim N(0,1)$. □

I have now shown two basic lemmas that will be used to show a very important property about Wiener processes: *Quadratic Variation*. This is a property that will allow for a well posed definition of the integral, because without this, the Riemann-like sums would not converge and I would not be able to define the stochastic integral properly.

2.2.2 Quadratic Variation

As stated previously, Wiener Processes have the unfortunate property that they are nowhere-differentiable. This, of course, makes it quite difficult to define things like stochastic differential equations when derivatives don't exist. It turns out, that although Wiener Processes lack differ-

entiability, they do have a rather important property that allows for a well posed notion of integrability: *Quadratic Variation*. This property allows one to construct a notion of integrability for stochastic processes because it causes the Riemann-like sums that will be used, to converge.

In order to continue, however, I will need to give a definition for what convergence means with respect to stochastic processes. I will start by giving a definition of an important space for stochastic processes.

Definition 2 ($L^2(\Omega)$ Space)**.** Given a set $\Omega \subseteq \mathbb{R}$, then define the set $L^2(\Omega)$ to be the set of all random variables X such that

$$\|X\| = \left(E\left[X^2\right]\right)^{\frac{1}{2}} < \infty.$$

This definition is important because it allows me to give a norm to random variables, and therefore, allows me to talk about what it means for random variables to *converge*. So, I will now define convergence for random variables.

Definition 3 (Random Variable Convergence)**.** Given a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$, this sequence is said to *converge* to a random variable Y in $L^2(\Omega)$ if

$$\|X_n - Y\| \to 0 \quad \text{as } n \to \infty.$$

This will be denoted as

$$X_n \xrightarrow{L^2} Y.$$

Now that I have defined what convergence for random variables will mean, I can discuss the quadratic variation property that Wiener Processes have. This property, as I have previously mentioned, is what allows for the stochastic integral to be well defined.

Lemma 3 (Quadratic Variation)**.** *Let* $[a, b]$ *be an interval in* $[0, \infty)$ *and suppose*

$$P^n := \{a = t_0^n < t_1^n < ... < t_{m_n}^n = b\}$$

are partitions of $[a, b]$*, with* $|P^n| \to 0$ *as* $n \to \infty$*. Also, let*

$$\tau_k = \lambda t_{k+1}^n + (1 - \lambda)\, t_k^n.$$

Then

$$\sum_{k=0}^{m_n - 1} \left[W(\tau_k) - W(t_k^n)\right]^2 \xrightarrow{L^2} \lambda(b - a).$$

Here, this lemma says that the sum of the *square* of the differences converges. Notice that this does not say anything about the sum of the differences. This quadratic variation property is slightly weaker of a property, but it is going to be very useful.

I will show the proof here, which comes from [15].

Proof. First, let $Q_n := \sum_{k=0}^{m_n - 1} \left(W(\tau_k) - W(t_k^n)\right)^2$. Instead of showing that Q_n converges to $\lambda(b - a)$, I will prove the quadratic variation property by showing that the difference between Q_n and $\lambda(b - a)$ converges to 0. In other words, I will show that

$$Q_n - \lambda(b - a) \xrightarrow{L^2} 0.$$

Notice that, I can rewrite the above difference as follows:

$$Q_n - \lambda(b - a) = \sum_{k=0}^{m_n - 1} \left[\left(W(\tau_k) - W(t_k^n)\right)^2 - \lambda(t_{k+1}^n - t_k^n)\right].$$

Recall from Definition 3 that in order to show that this converges to 0, I must show

$$E\left[\left(Q_n - \lambda(b - a)\right)^2\right] = 0.$$

Therefore, I will evaluate exactly this expectation. So, I have

$$E\left[(Q_n - \lambda(b-a))^2\right] =$$
$$\sum_{k=0}^{m_n-1}\sum_{j=0}^{m_n-1} E\left[\left(\left[W\left(\tau_k\right) - W\left(t_k^n\right)\right]^2 - \lambda\left[t_{k+1}^n - t_k^n\right]\right)\left(\left[W\left(\tau_j\right) - W\left(t_j^n\right)\right]^2 - \lambda\left[t_{j+1}^n - t_j^n\right]\right)\right]$$

This is a somewhat complicated expression, but it simplifies quite a bit because of the *independent increments* property of Wiener Processes. Notice that when $k \neq j$, the intervals $[t_k, \tau_k]$ and $[t_j, \tau_j]$ are non-overlapping. This follows directly from the fact that $[t_k, t_{k+1}] \cap [t_k, t_{k+1}] = \emptyset$ because of the construction of the partitions. So, because these intervals are non-overlapping, then I can break the expectation of the product to be the product of the expectation. In this case, when $k \neq j$, the term inside of the double sum becomes:

$$E\Big[\left(W(\tau_k) - W(t_k^n)\right)^2 - \lambda(t_{k+1}^n - t_k^n)\Big]E\left[\left(W(\tau_j) - W(t_j^n)\right)^2 - \lambda(t_{j+1}^n - t_j^n)\right]$$

Then, because $(W(t_s^n) - W(t_t^n)) \sim N(0, s-t)$, this product is 0. This means that the sum is only non-zero when $k = j$. So, I can simply replace the double sum with a single sum. This simplifies the equality greatly, and I am left with

$$E\left[(Q_n - \lambda(b-a))^2\right] = \sum_{k=0}^{m_n-1} E\left[\left(\left(W(\tau_k) - W(t_k^n)\right)^2 - \lambda(t_{k+1}^n - t_k^n)\right)^2\right].$$

At this point, I perform a clever trick of factoring out a $\lambda(t_{k+1}^n - t_k^n)$ from each term in the expectation term. This will result in

$$\begin{aligned} E\left[(Q_n - \lambda(b-a))^2\right] &= \sum_{k=0}^{m_n-1} E\left[\left(\left[\left(\frac{W(\tau_k) - W(t_k^n)}{\sqrt{\lambda(t_{k+1}^n - t_k^n)}}\right)^2 - 1\right]\left[\lambda(t_{k+1}^n - t_k^n)\right]\right)^2\right] \\ &= \sum_{k=0}^{m_n-1} E\left[\left(Y_k^2 - 1\right)^2\left(\lambda(t_{k+1}^n - t_k^n)\right)^2\right], \end{aligned}$$

where

$$Y_k = Y_k^n := \frac{W(\tau_k) - W(t_k^n)}{\sqrt{\lambda\left(t_{k+1}^n - t_k^n\right)}} \sim N(0,1).$$

This follows from Lemma 1. Continuing the proof, I see that

$$\begin{aligned} E\left[(Q_n - \lambda(b-a))^2\right] &= \sum_{k=0}^{m_n-1} E\left[\left(Y_k^2 - 1\right)^2 \left(\lambda(t_{k+1}^n - t_k^n)\right)^2\right] \\ &= \sum_{k=0}^{m_n-1} \left(\lambda(t_{k+1}^n - t_k^n)\right)^2 E\left[\left(Y_k^2 - 1\right)^2\right] \\ &= \sum_{k=0}^{m_n-1} 2\left(\lambda(t_{k+1}^n - t_k^n)\right)^2, \end{aligned}$$

where this last equality follows from Lemma 2. Finishing the proof, I see that

$$\begin{aligned} E\left[(Q_n - \lambda(b-a))^2\right] &= \sum_{k=0}^{m_n-1} 2\left(\lambda(t_{k+1}^n - t_k^n)\right)^2 \\ &= 2\lambda^2 \sum_{k=0}^{m_n-1} \left(t_{k+1}^n - t_k^n\right)^2 \\ &\leq 2\lambda^2 |P^n| \sum_{k=0}^{m_n-1} \left(t_{k+1}^n - t_k^n\right) \\ &= 2\lambda^2 |P^n|(b-a) \to 0 \text{ as } n \to \infty \text{ by assumption.} \end{aligned}$$

□

At this point, I've shown some basic stochastic lemmas and, most importantly, I've shown that Wiener Processes have the useful property of quadratic variation. As stated before, this is what will allow me to define what it means to integrate *with respect to* a random variable (namely the Wiener Process).

First, it will be helpful to try and integrate a much simpler random process

$$\int_0^T W dW, \tag{2.2}$$

instead of trying to integrate the very general random process given by G in Equation (2.1).

2.2.3 Integrate Wiener Process

Even though I could give a definition of the integral in Equation (2.1) at this point, it will be more helpful to discuss the easier integral

$$\int_0^T W dW,$$

as given in Equation (2.2).

This integral will be considered first because the Wiener Process is a much simpler stochastic process than the general stochastic process G. The idea will be to understand this simple integral, and then extend the ideas to the general integral.

Lemma 4. *Suppose P^n is a partition of the interval $[0, T]$ such that $|P^n| \to 0$ as $n \to \infty$. Suppose also that*

$$\tau_k = \lambda t_{k+1}^n + (1 - \lambda) t_k^n.$$

Then

$$\lim_{n\to\infty} \sum_{k=0}^{m_n - 1} W(\tau_k) \left(W(t_{k+1}^n) - W(t_k^n) \right) = \frac{1}{2} W(T)^2 + \left(\lambda - \frac{1}{2} \right) T.$$

In this case, I will define

$$\int_0^T W dW := \lim_{n\to\infty} \sum_{k=0}^{m_n - 1} W(\tau_k) \left(W(t_{k+1}^n) - W(t_k^n) \right).$$

Proof. This proof will rely mostly on Lemma 3 and comes largely from [15]. Here, I will finally use the quadratic variation property that I have been talking about so much. So, I will first start out by rewriting $W(\tau_k)$ as the following:

$$W(\tau_k) = [W(\tau_k) - W(t_k^n)] + \frac{1}{2} \left[W(t_{k+1}^n) + W(t_k^n) \right] - \frac{1}{2} \left[W(t_{k+1}^n) - W(t_k^n) \right]. \tag{2.3}$$

This may seem not seem intuitive, but it allows me to rewrite the infinite sum as several *quadratic variation*-looking sums. So, using 2.3 to rewrite $W(\tau_k)$, I get

$$\begin{aligned}
&\sum_{k=0}^{m_n-1}\left(\left[W(\tau_k)-W(t_k^n)\right]+\frac{1}{2}\left[W(t_{k+1}^n)+W(t_k^n)\right]-\frac{1}{2}\left[W(t_{k+1}^n)-W(t_k^n)\right]\right)\\
&\qquad\times\left(W(t_{k+1}^n)-W(t_k^n)\right)\\
&=\frac{1}{2}\sum_{k=0}^{m_n-1}W(t_{k+1}^n)^2-W(t_k^n)^2-\frac{1}{2}\sum_{k=0}^{m_n-1}\left(W(t_{k+1}^n)-W(t_k^n)\right)^2\\
&\qquad+\sum_{k=0}^{m_n-1}\left(W(\tau_k)-W(t_k^n)\right)\left(W(t_{k+1}^n)-W(t_k^n)\right)\\
&=\frac{1}{2}\sum_{k=0}^{m_n-1}W(t_{k+1}^n)^2-W(t_k^n)^2-\frac{1}{2}\sum_{k=0}^{m_n-1}\left(W(t_{k+1}^n)-W(t_k^n)\right)^2\\
&\qquad+\sum_{k=0}^{m_n-1}\left(W(\tau_k)-W(t_k^n)\right)\left(W(t_{k+1}^n)-W(\tau_k)+W(\tau_k)-W(t_k+1)\right).
\end{aligned}$$

At this point, it is helpful to go ahead and clean up a few terms. The first term in the above sum will telescope to give:

$$\begin{aligned}
\frac{1}{2}\sum_{k=0}^{m_n-1}W(t_{k+1}^n)^2-W(t_k^n)^2&=\frac{1}{2}\left[W(T)^2-W(0)^2\right]\\
&=\frac{1}{2}W(T)^2\qquad(W(0)=0\text{ a.s}).
\end{aligned}\tag{2.4}$$

Notice that this equality will hold even when this sum is passed to the limits.

The second term, $-\frac{1}{2}\sum_{k=0}^{m_n-1}\left(W(t_{k+1}^n)-W(t_k^n)\right)^2$, will be evaluated using Lemma 3. Here, I will use $\lambda=1$. So, passing this sum to the limits, and using Lemma 3, I conclude that

$$-\frac{1}{2}\sum_{k=0}^{m_n-1}\left(W(t_{k+1}^n)-W(t_k^n)\right)^2\xrightarrow{L^2}-\frac{1}{2}T.\tag{2.5}$$

The only term left to handle then, is the third term. This is a slightly trickier term but it is very close to being of the quadratic variation form. What I will do here, then, is to slightly rewrite what

is inside of the sum. Doing this, I get

$$\begin{aligned}
&\sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)\left(W(t_{k+1}^n) - W(t_k^n)\right) \\
&= \sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)\left(W(t_{k+1}^n) - W(\tau_k) + W(\tau_k) - W(t_k^n)\right) \\
&= \sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)\left(W(\tau_k) - W(t_k^n)\right) + \sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)\left(W(t_{k+1}^n) - W(\tau_k)\right) \\
&= \sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)^2 \qquad (2.6) \\
&\quad + \sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)\left(W(t_{k+1}^n) - W(\tau_k)\right). \qquad (2.7)
\end{aligned}$$

This simplification splits the sum into two easier sums. The sum (2.6), when passed to the limit, results in

$$\sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)^2 \xrightarrow{L^2} \lambda T \tag{2.8}$$

by Lemma 3 directly. The only thing left to show then, is that the sum (2.7) converges to 0. This proof is actually very close to the proof of Lemma 3.

I want to know what the expectation of the square of this term is. In other words, I would like to know what

$$E\left[\left(\sum_{k=0}^{m_n-1} \left(W(\tau_k) - W(t_k^n)\right)\left(W(t_{k+1}^n) - W(\tau_k)\right)\right)^2\right] \tag{2.9}$$

converges to, when passed to the limit.

I will start by expanding the square to rewrite Equation (2.9) as

$$E\left[\sum_{k=0}^{m_n-1}\sum_{j=0}^{m_n-1} \left[W(t_{k+1}) - W(\tau_k)\right]\left[W(\tau_k) - W(t_k)\right]\left[W(t_{j+1}) - W(\tau_j)\right]\left[W(\tau_j) - W(t_j)\right]\right]. \tag{2.10}$$

When $k \neq j$, all four of these products are independent because of the independent increments property of Wiener Processes. Thus, Equation 2.10 results in

$$\sum_{k=0}^{m_n-1} \sum_{j=0}^{m_n-1} E\left[\left(W(\tau_k) - W(t_k^n)\right)\right] E\left[\left(W(t_{k+1}^n) - W(\tau_k)\right)\right] \times E\left[\left(W(\tau_j) - W(t_j^n)\right)\right] E\left[\left(W(t_{j+1}^n) - W(\tau_j)\right)\right].$$

Notice that, by the second property of Wiener Processes, each of these terms in the product is 0. Thus, this entire sum is simply the sum of zeros.

So, Equation (2.10) can be simplified to combine the summing indices to result in

$$E\left[\sum_{k=0}^{m_n-1} \left([W(t_{k+1}) - W(\tau_k)]\left[W(\tau_k) - W(t_k)\right]\right)^2\right]. \tag{2.11}$$

Because the intervals $t_k^n \leq \tau_k \leq t_{k+1}^n$, then the random variables $(W(\tau_k) - W(t_k^n))$ and $\left(W(t_{k+1}^n) - W(\tau_k)\right)$ are independent by the independent increments property of Wiener Processes. This means that I can split the expectation one last time to result in

$$\sum_{k=0}^{m_n-1} E\left[\left[W(t_{k+1}) - W(\tau_k)\right]^2\right] \times E\left[\left[W(\tau_k) - W(t_k)\right]^2\right]. \tag{2.12}$$

Then, by properties of Wiener processes, this simplifies to

$$\begin{aligned}
\sum_{k=0}^{m_n-1} \left(t_{k+1}^n - \tau_k\right)\left(\tau_k - t_k^n\right) &= \sum_{k=0}^{m_n-1} (1-\lambda)\left(t_{k+1}^n - t_k^n\right)\lambda\left(t_{k+1}^n - t_k^n\right) \\
&= \sum_{k=0}^{m_n-1} \lambda(1-\lambda)\left(t_{k+1}^n - t_k^n\right)^2 \\
&\leq \lambda(1-\lambda)\,|P^n|T \to 0 \qquad \text{as } n \to \infty.
\end{aligned}$$

So, since sum (2.7) converges to zero when passed to the limit, I can deduce that

$$\lim_{n\to\infty} \sum_{k=0}^{m_n-1} W(\tau_k)\left(W(t_{k+1}^n) - W(t_k^n)\right) = \frac{1}{2}W(T)^2 + \left(\lambda - \frac{1}{2}\right)T \qquad a.s.$$

□

Notice that this integral will converge to different values depending on the "evaluation" point τ_k that is chosen in each interval. This is very much unlike the "normal" integral, where the evaluation point doesn't matter. So, a choice must be made on the evaluation point. This is actually a topic of debate, and has led to what is known as the *Itô-Stratonovich Controversy* [20]. For this thesis, however, I will be making the decision to use what is called the *Itô Definition*. This corresponds to making the choice of setting $\lambda = 0$. Thus, with this choice of the evaluation point, I can finally define this integral completely as

$$\int_0^T W dW := \frac{1}{2}W(T)^2 - \frac{1}{2}T. \tag{2.13}$$

2.2.4 General Stochastic Integral

With the integral of the very specific stochastic integral in Equation (2.2) defined, I can now define the general stochastic integral given in Equation (2.1). Because this chapter of the thesis is meant to give a very brief overview of stochastic calculus, then I will not be going into the full details of the proof of the general stochastic integral. There are some subtleties in the proof that go beyond the scope of this thesis, but for more information, see [15]. I will, however, give a brief overview of the logic but leave the details to the reader. The ideas outlined here come from [15].

The idea is that I will consider a simpler stochastic process, called a *step process*. I will define the stochastic integral for this simpler process, and then approximate the more general stochastic process with a series of these step processes.

First, I will give a definition of an important space, $L^2(0,T)$.

Definition 4. $L^2(0,T)$ is defined to be the space of all real-valued, progressively measureable stochastic processes $G(\cdot)$ such that

$$E\left[\int_0^T G^2 dt\right] < \infty.$$

Similarly, the space $L^1(0,T)$ is the space of all real-valued, progressively measureable stochastic processes $G(\cdot)$ such that

$$E\left[\int_0^T |G| dt\right] < \infty.$$

These spaces of stochastic processes are essentially the "well-behaved" stochastic processes that I mentioned at the beginning of this chapter. I did not give a definition of *progressively measureable*, becuase it goes outside the scope of this thesis, but think of it as only depending on the history of Wiener Processes.

It will also be useful to define what a *step process* is.

Definition 5. A process $G \in \mathbb{L}^2(0,T)$ is called a *step process* if there exists a partition $P = \{0 = t_0 < t_1 < ... < t_m = T\}$ such that

$$G(t) = G_k \qquad \text{for } t_k \leq t < t_k + 1 \quad (k = 0, ..., m-1).$$

This just means that there is some partition where the value of G is given by the value at the beginning of each interval. The *beginning* of each interval is important, because when I defined the integral of a Wiener Process with respect to a Wiener Process in Equation (2.2), I evaluated the process at the beginning of each interval. The fact that my step process only depends on the value at the beginning of each interval will allow me to give a consistent definition to the stochastic integral.

I went through the work of defining the stochastic integral of a Wiener Process because it showed me that for these stochastic integrals, I will probably have to care about the evaluation point in each interval. Then, because I chose to evaluate my Wiener Process at the beginning

of each interval, I should probably make a similar decision in all other cases too, in order to be consistent.

So, with the perspective from the stochastic integral of the Wiener Process, I will define the stochastic integral of this step process to be

Definition 6. Let $G \in L^2(0,T)$ be a step process. Then,

$$\int_0^T GdW \coloneqq \sum_{k=0}^{m_n-1} G_k \left(W(t_{k+1}) - W(t_k)\right).$$

As mentioned previously, because of the choice of evaluating the step process at the beginning of each interval, this integral is really called the *Itô Integral*, although in this thesis I will sometimes refer to this as simply the stochastic integral. It is also good to note that this integral is a random variable. This is somewhat hard to remember because of being used to integrals resulting in a *number*, but in this case, the integral results in a random variable.

This integral, as defined in Definition 6, has many of the same properties that an integral would be expected to have, such as linearity, but I will not be discussing them here. Refer to [15] for more information on these properties.

The only thing left to talk about, is the general stochastic (or Itô) integral mentioned in Equation (2.1). Again, I will not be going into the details of the full proof of this here, but the idea is that I can approximate any stochastic process $G \in L^2(0,T)$ by a series of step processes. I will then define the integral below.

Definition 7. Let $G \in L^2(0,T)$. Then let $\{G^n\}$ be a series of step processes that approximates G. Then define

$$\int_0^T GdW \coloneqq \lim_{n\to\infty} \int_0^T G_n dW.$$

It turns out that this integral is well defined, and so this definition gives a well posed definition of the stochastic integral. For details on the proof of the well-posedness of this definition, refer to [15].

So, finally I have a definition for what it means to integrate a stochastic process with respect to a Wiener Process. I started this section by talking about some basic lemmas that were important in the proof of a very important property of Wiener Processes: Qaudratic Variation. Then, with this property understood, I looked at how I could define the stochastic integral of a Wiener Process. This lead me to understand that a choice needed to be made when it comes to the evaluation point of the Wiener Process. I chose to make the decision to evaluate the Wiener Process at the beginning of each interval, which turns out to be the definition of the Itô integral. Then, with this understanding of where to evaluate my stochastic process, I extended this insight to define the general stochastic integral. This integral has been defined through a Riemann-like argument, and with this definition, I will finally be able to talk about the most important part of this stochastic background: the *stochastic differential*.

2.3 Stochastic Differential Equations

Recall that the Black-Scholes Model is a PDE, but the underlying stocks will be modeled as stochastic processes. This means that at some point, I will need to understand how stochastic processes fit into differential equations. Put another way, I will need to understand the topic of *Stochastic Differential Equations* to some degree.

Regular differential equations, which take the form of

$$\begin{cases} \frac{dx}{dt} = f(t), \\ x(0) = x_0, \end{cases} \tag{2.14}$$

are well understood to mean

> *What is the function $x(t)$ such that its derivative with respect to time is equal to $f(t)$, and $x(0) = x_0$?*

It seems reasonable to extend this definition to stochastic processes by adding effects of randomness. This could be done by letting $X(t)$ be a stochastic process, and looking at its differential

equation given by

$$\begin{cases} dX(t) = F(t)dt + G(t)dW(t), \\ X(0) = X_0, \end{cases} \tag{2.15}$$

where $F(t)$ and $G(t)$ are stochastic processes. I will show later that they will need to satisfy some conditions, but for now, I will simply think of them as stochastic processes.

The statement given in Equation (2.15), cannot be taken to mean the same thing that the differential equation in Equation (2.14) means. This is because of the nowhere-differentiability problem of Wiener Processes. It turns out, that because of this property, a derivative of the stochastic process $X(t)$ just doesn't exist. This means then, that I cannot take the statement in Equation (2.15) to mean the same thing that the statement in Equation (2.14) means. I will have to come up with another understanding of what Equation (2.15) means.

This is where all of the work from Section 2.2 will be useful. Since I cannot understand the **Stochastic Differential Equation (SDE)** in Equation (2.15) in terms of derivatives, then maybe I can understand it in terms of *integrals*. This takes inspiration from the Second Fundamental Theorem of Calculus, which (loosely) states that a function is equal to the integral of its derivative. I will take that inspiration and make sense of the SDE in Equation (2.15) through stochastic integration.

Definition 8. Let $X(t)$ be a stochastic process that satisfies

$$X(T) = X_0 + \int_0^T F(t)dt + \int_0^T G(t)dW(t)$$

for $F \in L^1(0,T)$, $G \in L^2(0,T)$, and $X(0) = X_0$. Then, the *differential* of $X(t)$, denoted $dX(t)$ is given by

$$dX(t) = F(t)dt + G(t)dW(t).$$

One important thing to note here is that the terms dX, dt and dW don't really *mean* anything more than simply being a mathematical abbreviation.

With this definition, there is now meaning to what the differential of a stochastic process is and how to interpret what a the related SDE means. To be explicit, when given the SDE given in Equation (2.15), this is to be interpreted as finding the stochastic process $X(t)$ given by

$$X(t) = X_0 + \int_0^T F ds + \int_0^T G dW.$$

This is extremely interesting because it *defines* the differential in terms of the integral. This is in contrast to "ordinary" calculus where the differential is defined in terms of the derivative.

The goal of this section is to now to get an understanding of the most important lemma in this chapter: Itô's Lemma. This is a fairly complicated and profound lemma that essentially extends the chain rule to stochastic calculus. Therefore, I will first provide some basic stochastic calculus lemmas that will ultimately lead to a proof of the famous Itô's Lemma.

2.3.1 Basic Stochastic Differentials

This section aims at describing some basic stochastic differentials in order to gain an understanding of these somewhat strange differentials. Because these SDEs are given meaning through the context of integrals, they are somewhat difficult to understand. Therefore, I find it helpful to first look at some basic differentials that will also be useful in some proofs in later sections.

The first of these basic differentials I will outline in this thesis is $d(W^2)$. The reason for this one is because I have already done much of the work in Section 2.2.3. Recall that at the end of this section, I came to the conclusion that

$$\int_0^T W dW = \frac{1}{2}\left(W(T)^2 - T\right).$$

Notice that I can rearrange some of these terms to get the following expression

$$W(T)^2 = T + 2\int_0^T W dW. \tag{2.16}$$

Here, I notice that I can rewrite the term T to be an integral with respect to time. In particular, this can be rewritten as

$$T = \int_0^T dt.$$

Substituting this into Equation (2.16), I get

$$W(T)^2 = \int_0^T dt + 2\int_0^T W dW.$$

This looks just like the expression in Definition 8 with

- W(0) = 0,
- F(t) = 1 and
- G(t) = 2W(t).

Therefore, by definition, I say

$$d(W^2) = dt + 2W dW. \tag{2.17}$$

One thing to note at this point is that this looks *almost* like what I would expect this differential to be if I were using "normal" calculus. If so, I would expect this differential to be $2W dW$, but here I have the extra dt term. This term is sometimes called the Itô Correction Term [15].

In this differential, I actually started with the integral to expose the differential. This will be a common theme when showing what stochastic differentials equal.

The next basic differential I would like to discuss is $d(tW)$. If I were using "regular" calculus, I would guess that this differential would be

$$W dt + t dW$$

by the product rule. It turns out that this happens to be what this differential equals, but I will show this below.

Lemma 5. *The following is true:*

$$d(tW) = Wdt + tdW.$$

Proof. This proof is taken from [15]. I will add a few of my own notes and understanding to this proof, but it is largely taken from this source.

To begin, I will note that I am trying to show

$$TW(T) = \int_0^T Wds + \int_0^T tdW.$$

It is helpful to note that, by definition, I have

$$\int_0^T Wds = \lim_{n\to\infty} \sum_{k=0}^{m_n-1} W(t_k^n)(t_{k+1}^n - t_k^n).$$

Because Wiener Processes have continuous sample paths, then this can be interpreted as a regular Riemann integral. Therefore, I am not restricted to evaluating my stochastic process at the left endpoint of each interval. I am allowed to evaluate my stochastic process at any point in each interval. So, I will take the liberty of evaluating my stochastic process at the *right* endpoint of each interval. The reason for this will become obvious later. Therefore, I actually have

$$\int_0^T Wds = \lim_{n\to\infty} \sum_{k=0}^{m_n-1} W(t_{k+1}^n)(t_{k+1}^n - t_k^n).$$

Next, I will tackle the second integral. Again, by definition I have:

$$\int_0^T tdW = \lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_k^n(W(t_{k+1}^n) - W(t_k^n)).$$

Now that I have expanded each of these integrals into their Riemann-like sums, I am going to combine them and see what happens. So, I have

$$\begin{aligned}\int_0^T W ds + \int_0^T t dW &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} W(t_{k+1}^n)(t_{k+1}^n - t_k^n) \; + \lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_k^n (W(t_{k+1}^n) - W(t_k^n)) \\ &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_{k+1}^n W(t_{k+1}^n) - t_k^n W(t_{k+1}^n) + t_k^n W(t_{k+1}^n) - t_k^n W(t_k^n) \\ &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_{k+1}^n W(t_{k+1}^n) - t_k^n W(t_k^n).\end{aligned}$$

This last line is because the middle two terms cancelled out. Notice that at this point, this sum forms a telescoping sum. When this sum is simplified, I arrive at the following:

$$\int_0^T W ds + \int_0^T t dW = TW(T) - 0W(0) = TW(T).$$

So, I have indeed shown that

$$TW(T) = \int_0^T W ds + \int_0^T t dW,$$

which means that I am done, and therefore have shown that

$$d(tW) = W dt + t dW.$$

□

This is interesting because in this case I don't have an Itô Correction Term. The differential is exactly what I would have expected if I were using "regular" calculus. The reason for this is really because the term t is deterministic. Because of this deterministic term, the Itô Correction Term becomes zero, and the result is what "regular" calculus would yield.

At this point I have given examples of a few simple stochastic differentials, along with their proofs. Unfortunately, these have all been *one dimension*, or have all only had a single Wiener

Process involved. Because this thesis focuses on the Muti-Asset Black-Scholes Model, then it seems reasonable that I will want to extend some of these ideas to multiple dimensions, or multiple Wiener Processes.

In the next section, I will discuss some basic stochastic differentials in higher dimensions to get an understanding of how the things discussed in this section generalize to higher dimensions.

2.3.2 Basic Multi-Dimensional Stochastic Differentials

In this section, I will extend some of the concepts discussed in Section 2.3.1. This will include a generalization of Lemma 5 as well as a multi-dimensional version of the differential of the product of two *independent* Wiener Processes.

First, I will start with a generalization of Lemma 5.

Lemma 6. *Suppose* $\{W_i\}_{i\in\mathbb{N}}$ *are independent Wiener Processes. Then*

$$d(tW_i) = W_i dt + t dW_i, \quad \forall i \in \mathbb{N}.$$

The proof of this lemma will be omitted, because it is an easy extension of the proof of Lemma 5.

Before I can introduce the more interesting lemma about the differential of a product of independent Wiener Processes I need to introduce a lemma. This is a lemma discussing the distribution of a normalized sum of independent Wiener Processes.

Lemma 7. *Suppose that* W_i *and* W_j *are two independent Wiener Processes. Then, define*

$$X(t) \coloneqq \frac{W_i(t) + W_j(t)}{\sqrt{2}}.$$

Then, $X(t)$ *is a Wiener Process.*

Proof. This proof is outlined in [15], but I have added the details of the proof here.

To prove this, there are three things that need to be shown

1. $X(0) = 0$ a.s.,

2. $X(t) - X(s) \sim N(0, t-s), \quad \forall 0 \leq s \leq t$,

3. $\forall 0 \leq t_1 \leq ... \leq t_n$, the random variables $X(t_1)$, $X(t_2) - X(t_1)$, ..., $X(t_n) - X(t_{n-1})$ are independent.

The first part is somewhat trivial. Because both $W_i(0) = W_j(0) = 0$ almost surely, and because $X(0)$ is simply the sum of these two random variables, then $X(0) = 0$ almost surely as well.

Before I can show the second part, it is necessary to note that $X(t) \sim N(0, t)$. This mainly comes from the fact that $X(t)$ is the difference of two independent normally distributed random variables. This fact immediately leads to the conclusion that $X(t)$ is normally distributed. To show that the mean is zero, I can employ the fact that the expectation behaves linearly. So,

$$\begin{aligned} E[X] &= \frac{E[W_i] + E[W_j]}{\sqrt{2}} \\ &= \frac{0+0}{\sqrt{2}} \\ &= 0. \end{aligned}$$

The variance also behaves fairly nicely as well. The only catch is that the coefficient needs to be squared. Therefore, I have

$$\begin{aligned} \text{var}(X) &= \text{var}\left(\frac{W_i + W_j}{\sqrt{2}}\right) \\ &= \left(\frac{1}{\sqrt{2}}\right)^2 (\text{var}(W_i) + \text{var}(W_j)) \\ &= \frac{1}{2}(t+t) \\ &= t. \end{aligned}$$

Therefore, $X(t) \sim N(0, t)$.

Now, to prove the second part of the definition for Wiener Processes is true, I will need to show that $X(t) - X(s) \sim N(0, t-s)$ for all $t \geq s$. In order to show this, I will look directly at the random variable $X(t) - X(s)$:

$$\begin{aligned} X(t) - X(s) &= \frac{W_i(t) + W_j(t)}{\sqrt{2}} - \frac{W_i(s) + W_j(s)}{\sqrt{2}} \\ &= \frac{W_i(t) - W_i(s)}{\sqrt{2}} + \frac{W_j(t) - W_j(s)}{\sqrt{2}}. \end{aligned} \tag{2.17}$$

At this point, I will need to show that the two random variables in Equation (2.17) are independent. The reason for this is because I need these to be independent in order to show that the $X(t) - X(s)$ is normally distributed. So, to show the independence, I will show that the expectation of the product is the product of the expectations. I have:

$$\begin{aligned} &E\left[\frac{W_i(t) - W_i(s)}{\sqrt{2}} \cdot \frac{W_j(t) - W_j(s)}{\sqrt{2}}\right] \\ &= \frac{1}{2} E\left[W_i(t)W_j(t) - W_i(t)W_j(s) - W)i(s)W_j(t) + W_i(s)W_j(s)\right] \\ &= \frac{1}{2}\left(E\left[W_i(t)W_j(t)\right] - E\left[W_i(t)W_j(s)\right] - E\left[W_i(s)W_j(t)\right] + E\left[W_i(s)W_j(s)\right]\right). \end{aligned}$$

Because $W_i(t)$ is independent of $W_j(t)$ for all $t \geq 0$ by assumption, then the arguments for each of the expectations above are independent. This results in

$$\begin{aligned} &\frac{1}{2}\left(E\left[W_i(t)W_j(t)\right] - E\left[W_i(t)W_j(s)\right] - E\left[W_i(s)W_j(t)\right] + E\left[W_i(s)W_j(s)\right]\right) \\ &= \frac{1}{2}\left(E\left[W_i(t)\right]E\left[W_j(t)\right] - E\left[W_i(t)\right]E\left[W_j(s)\right] - E\left[W_i(s)\right]E\left[W_j(t)\right] + E\left[W_i(s)\right]E\left[W_j(s)\right]\right) \\ &= \frac{1}{2}\left(E\left[W_i(t)\right] - E\left[W_i(s)\right]\right) \cdot \left(E\left[W_j(t)\right] - E\left[W_j(s)\right]\right) \\ &= \frac{1}{2} E\left[W_i(t) - W_i(s)\right] \cdot E\left[W_j(t) - W_j(s)\right] \\ &= E\left[\frac{W_i(t) - W_i(s)}{\sqrt{2}}\right] \cdot E\left[\frac{W_j(t) - W_j(s)}{\sqrt{2}}\right]. \end{aligned}$$

This then, shows that the random variables in Equation (2.17) are independent because the expectations of their products is the product of their expectations. This is to be expected, however, because both W_i and W_j are independent. Now, because of the independence, I can say that $X(t) - X(s)$ is normally distributed because it is the sum of two independent normally distributed random variables. The expectation and variance, are more easily shown. For the expectation,

$$\begin{aligned}
&E\left[\frac{W_i(t) - W_i(s)}{\sqrt{2}} + \frac{W_j(t) - W_j(s)}{\sqrt{2}}\right] \\
&= E\left[\frac{W_i(t) - W_i(s)}{\sqrt{2}}\right] + E\left[\frac{W_j(t) - W_j(s)}{\sqrt{2}}\right] \\
&= 0.
\end{aligned}$$

The last line uses the definition of Wiener Processes. For the variance,

$$\begin{aligned}
&\text{var}\left(\frac{W_i(t) - W_i(s)}{\sqrt{2}} + \frac{W_j(t) - W_j(s)}{\sqrt{2}}\right) \\
&= \frac{1}{2}\left(\text{var}\left(W_i(t) - W_i(s)\right) + \text{var}\left(W_j(t) - W_j(s)\right)\right) \\
&= \frac{1}{2}\left((t-s) + (t-s)\right) \\
&= t - s.
\end{aligned}$$

Therefore, $X(t) - X(s) \sim N(0, t-s)$.

The last step is to show the independent increments property. The exact steps of this will not be shown in this thesis because it follows similar steps to what is shown above. To summarize, however, it is done through an induction process. At each step, one will need to show that the expectation of the product is the product of the expectations. Then, when expanding the terms, one will need to use the properties of the Wiener Process and the assumption that W_i and W_j are independent.

With all three properties shown, I can conclude that $X(t)$ is a Wiener Process.

□

Now to the more interesting lemma. This lemma talks about the differential of a product of independent Wiener Processes. This will be useful in the next section when I prove the product rule for stochastic differentials.

Lemma 8. *Suppose* $\{W_i\}_{i\in\mathbb{N}}$ *are independent Wiener Processes. Then*

$$d(W_iW_j) = W_i dW_j + W_j dW_i, \quad i \neq j.$$

Proof. This proof is also taken from [15], where again, I have added some of my own personal annotations and understanding.

I will start by considering the random variable

$$X(t) \coloneqq \frac{W_i(t) + W_j(t)}{\sqrt{2}}. \tag{2.18}$$

This is the same definition as given in Lemma 7. So, because of this and Equation (2.17), I have the following equalities

$$\begin{cases} d(X^2) & = 2XdX + dt \\ d(W_i^2) & = 2W_i dW_i + dt \\ d(W_j^2) & = 2W_j dW_j + dt. \end{cases}$$

Also, I will note that

$$W_iW_j = X^2 - \frac{1}{2}W_i^2 - \frac{1}{2}W_j^2.$$

Then,

$$d(W_iW_j) = d\left(X^2 - \frac{1}{2}W_i^2 - \frac{1}{2}W_j^2\right).$$

Since the differential operator is linear (because the integral is linear), then I have the following

$$\begin{aligned} d\left(X^2 - \frac{1}{2}W_i^2 - \frac{1}{2}W_j^2\right) &= d\left(X^2\right) - \frac{1}{2}\left(W_i^2\right) - \frac{1}{2}\left(W_j^2\right) \\ &= 2XdX + dt - W_i dW_i - \frac{1}{2}dt - W_j dW_j - \frac{1}{2}dt \\ &= 2XdX - W_i dW_i - W_j dW_j \\ &= (W_i + W_j)\cdot(dW_i + dW_j) - W_i dW_i - W_j dW_j \\ &= W_i dW_j + W_j dW_i \end{aligned}$$

□

This is interesting because, in this case, there is no Itô correction term. It turns out, that this is because the two Wiener Processes are independent.

So far, I have shown a few important stochastic differentials. These differentials were mostly differentials of products of simple processes. During the first few differentials, I showed the equality of the differentials through the definition, which took the path of using the stochastic integral. This was important because it showed the usefulness and necessity of having the stochastic integral before talking about the stochastic differential. Then, in the last differential, I took the path of looking at a related stochastic process and showing that this process, which I called $X(t)$ was itself a Wiener Process. This then allowed me to use some of the theorems that I had shown already to rewrite the differential into an easier form.

All of these differentials will be useful in the next section, where I talk about the product rule for stochastic differentials. I have spent quite a bit of time in this section describing the differential for products of simple processes, and now I will use these results to talk about the differential of a generalized product.

2.3.3 Product Rule

With all of the knowledge so far, I can finally talk about the product rule for stochastic differentials. This is useful in its right, but it will be particularly useful in proving the most important result of this chapter: Itô's Lemma.

In this section, and in the next section, I will be talking about the multi-dimensional version of these theorems. This is because my goal is to be able to use these theorems in the Basket-Case version of the Black-Scholes Model.

From this point onward, I will also be using what is called Einstein Notation [6]. This makes the calculations and notation quite a bit simpler, and is the main reason for adopting this convention from here on out.

One other notational thing to discuss is how I will be representing vectors and matrices. I have already chosen to represent random variables with a capitol letter, so it seems natural to represent a vector of random variables as a capital bold faced letter. This conflicts with the common notational choice to represent matrices with a capital bold faced letter, so I will make the following notational choice:

$$\begin{cases} \mathbf{a} - \mathbf{z} & \text{for vectors in } \mathbb{R}^n, \\ \mathbf{A} - \mathbf{M} & \text{for matrices and} \\ \mathbf{N} - \mathbf{Z} & \text{for vectors of random variables.} \end{cases}$$

With that out of the way, I will move on to the statement and proof of the product rule.

Lemma 9 (Product Rule)**.** *Suppose* $\left\{W^k\right\}_{0\leq k\leq n}$ *are i.i.d. Wiener Processes. Then, let*

$$\begin{cases} dX_1 = F_1 dt + G_1^k dW^k \\ dX_2 = F_2 dt + G_2^k dW^k. \end{cases}$$

for $F_i \in L^1(0,T)$ and $G_i^k \in L^2(0,T)$ for $i \in \{1,2\}$ and $k \in \{1,...,n\}$. Note that Einstein Notation is being used so that $G_1^k dW^k$ should be taken to mean $\sum_{k=1}^n G_1^k dW^k$. Then,

$$d\left(X_1X_2\right) = \nabla_X(X_1X_2)\cdot d\mathbf{X} + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(X_1X_2)dt, \tag{2.19}$$

where $\mathbf{G}_{ij} = G_i^j$, $\nabla_X(\cdot)$ represents the gradient operator and $\nabla_X^2(\cdot)$ represents the Hessian.

Proof. A one dimensional version of this proof was given in [15], but I have adapted this proof to multi-dimensions.

I will assume that $F_i(t)$ and $G_i^j(t)$ are step processes. For the sake of this thesis, I will only show this is true for step processes. This won't quite prove the lemma itself, but the remainder of the proof can be seen by approximating the general random processes with series of step processes to complete the proof.

I will start by partitioning my interval up into n sub-intervals, corresponding to the sub-intervals where F_i and G_i^j are constant.

Recall from the definition of the stochastic differential, that I need to show

$$X_1X_2 = \sum_{k=0}^{m_n-1}\left(\int_{t_k}^{t_{k+1}} \nabla_X(X_1X_2)\cdot d\mathbf{X} + \frac{1}{2}\int_{t_k}^{t_{k+1}} \left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(X_1X_2)dt.\right) \tag{2.20}$$

This equation is slightly different than the definition, but the difference is simply that I am integrating over each sub-interval and then adding up these integrals.

So, I will evaluate this expression and validate that it equals X_1X_2. It will be easier if I first rewrite the integrands using Einstein Notation.

Looking at the first integrand, I see

$$\begin{aligned}\nabla_X(X_1X_2)\cdot d\mathbf{X} &= \left(\nabla_X(X_1X_2)\right)_i \left(d\mathbf{X}\right)_i \\ &= \frac{\partial}{\partial X_i}(X_1X_2)\cdot dX_i.\end{aligned}$$

Since $i = 1, 2$, then this can easily be explicitly written out:

$$\nabla_X (X_1 X_2) \cdot d\mathbf{X} = X_2 dX_1 + X_1 dX_2. \tag{2.21}$$

The second integrand is somewhat more complicated, but when it is expanded, I have:

$$\begin{aligned} \left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2 (X_1 X_2)\, dt &= \left(\mathbf{G}\mathbf{G}^T\right)_{ij} \left(\nabla_X^2 (X_1 X_2)\right)_{ij} dt \\ &= \mathbf{G}_{ik}\mathbf{G}_{jk} \frac{\partial}{\partial X_i \partial X_j} (X_1 X_2)\, dt. \end{aligned} \tag{2.22}$$

This looks complicated, but it turns out that I have the following:

$$\begin{cases} \frac{\partial}{\partial X_i \partial X_j} (X_1 X_2) = 0 & , i = j \\ \frac{\partial}{\partial X_i \partial X_j} (X_1 X_2) = 1 & , i \neq j, \end{cases}$$

where $i, j \in \{1, 2\}$. This means that I can use the Kronecker Delta symbol to rewrite this term as the following:

$$\frac{\partial}{\partial X_i \partial X_j} (X_1 X_2) = 1 - \delta_{ij}.$$

Now, substituting this into (2.22), and using the symmetry of i and j, I have

$$\mathbf{G}_{ik}\mathbf{G}_{jk} \cdot \frac{\partial}{\partial X_i \partial X_j} (X_1 X_2)\, dt = 2 G_1^k G_2^k dt. \tag{2.23}$$

At this point, I am going to substitute both (2.21) and (2.23) into the right hand side of (2.20). I am also going to evaluate these integrals for a fixed k, and then do the summing at the end. This

is because the notation is easier to follow this way. So, I have:

$$\begin{aligned}
&\int_{t_k}^{t_{k+1}} X_2 dX_1 + \int_{t_k}^{t_{k+1}} X_1 dX_2 + \frac{1}{2}\int_{t_k}^{t_{k+1}} 2G_1^l G_2^l dt \\
&= \int_{t_k}^{t_{k+1}} \left(F_2 t + G_2^p W^p\right)\left(F_1 dt + G_1^q dW^q\right) + \int_{t_k}^{t_{k+1}} \left(F_1 t + G_1^s W^s\right)\left(F_2 dt + G_2^d dW^d\right) \\
&\quad + \int_{t_k}^{t_{k+1}} G_1^l G_2^l dt.
\end{aligned}$$

At this point, I will use the fact that each G_i^j is a step process. Because I chose my intervals in such a way that each G_i^j is constant, then G_i^j is not time dependent in these integrals, so they can be pulled out of the time integral above. Thus, I have:

$$\begin{aligned}
&\int_{t_k}^{t_{k+1}} \left(F_2 t + G_2^p W^p\right)\left(F_1 dt + G_1^q dW^q\right) + \int_{t_k}^{t_{k+1}} \left(F_1 t + G_1^s W^s\right)\left(F_2 dt + G_2^d dW^d\right) \\
&\quad + \int_{t_k}^{t_{k+1}} G_1^l G_2^l dt \\
&= \int_{t_k}^{t_{k+1}} \left(F_2 t + G_2^p W^p\right)\left(F_1 dt + G_1^q dW^q\right) + \int_{t_k}^{t_{k+1}} \left(F_1 t + G_1^s W^s\right)\left(F_2 dt + G_2^d dW^d\right) \\
&\quad + G_1^l G_2^l (t_{k+1} - t_k).
\end{aligned}$$

The next step here is to expand the other two integrals above. Distributing and collecting like terms,

$$\begin{aligned}
&\int_{t_k}^{t_{k+1}} \left(F_2 t + G_2^p W^p\right)\left(F_1 dt + G_1^q dW^q\right) + \int_{t_k}^{t_{k+1}} \left(F_1 t + G_1^s W^s\right)\left(F_2 dt + G_2^d dW^d\right) \\
&\quad + G_1^l G_2^l (t_{k+1} - t_k) \\
&= 2\int_{t_k}^{t_{k+1}} F_1 F_2 t dt \\
&\quad + \int_{t_k}^{t_{k+1}} F_2 G_1^q t dW^q + \int_{t_k}^{t_{k+1}} F_2 G_1^s W^s dt \\
&\quad + \int_{t_k}^{t_{k+1}} F_1 G_2^d t dW^d + \int_{t_k}^{t_{k+1}} F_1 G_2^p W^p dt \\
&\quad + \int_{t_k}^{t_{k+1}} G_2^p G_1^q W^p dW^q + \int_{t_k}^{t_{k+1}} G_1^s G_2^d W^s dW^d \\
&\quad + G_1^l G_2^l (t_{k+1} - t_k).
\end{aligned}$$

At this point, it is important to notice that many of the indices are "dummy" indices, and as long as I am careful, I can rename them. Therefore, I will do some renaming of the indices to allow for some of the terms to be grouped together.

$$\begin{aligned}
&2\int_{t_k}^{t_{k+1}} F_1 F_2 t dt \\
&\quad + F_2 G_1^q \left(\int_{t_k}^{t_{k+1}} t dW^q + \int_{t_k}^{t_{k+1}} W^q dt\right) + F_1 G_2^d \left(\int_{t_k}^{t_{k+1}} t dW^d + \int_{t_k}^{t_{k+1}} W^d dt\right) \\
&\quad + G_2^p G_1^s \left(\int_{t_k}^{t_{k+1}} W^p dW^s + \int_{t_k}^{t_{k+1}} W^s dW^p\right) + G_1^l G_2^l (t_{k+1} - t_k) \\
&= 2\int_{t_k}^{t_{k+1}} F_1 F_2 t dt \\
&\quad + \left(F_1 G_2^d + F_2 G_1^d\right)\left(\int_{t_k}^{t_{k+1}} t dW^d + \int_{t_k}^{t_{k+1}} W^d dt\right) \\
&\quad + G_2^p G_1^s \left(\int_{t_k}^{t_{k+1}} W^p dW^s + \int_{t_k}^{t_{k+1}} W^s dW^p\right) + G_1^l G_2^l (t_{k+1} - t_k).
\end{aligned}$$

At this point, I have many integrals that look familiar. In fact, I derived many of them in Section 2.3.2. Substituting the results from Section 2.3.2, and using the fact that F_i is constant on the interval of integration, I have

$$\begin{aligned}
&F_1F_2(t_{k+1}^2 - t_k^2) \\
&\quad + \left(F_1G_2^d + F_2G_1^d\right)\left(t_{k+1}W^d(t_{k+1}) - t_kW^d(t_k)\right) \\
&\quad + G_2^pG_1^s\left(W^p(t_{k+1})W^s(t_{k+1}) - W^p(t_k)W^s(t_k)\right)(1-\delta_{ps}) \\
&\quad + G_2^pG_1^p\left(W^p(t_{k+1}) - t_{k+1} - W^p(t_k) + t_k\right) \\
&\quad + G_1^lG_2^l(t_{k+1} - t_k).
\end{aligned}$$

This may not look like the result that I am looking for, but I still need to sum over k. When I do this, many of the terms above form a telescoping series, and thus cancel out nicely. This will result in

$$\begin{aligned}
&\sum_{k=0}^{m_n-1}\left(\int_{t_k}^{t_{k+1}}\mathbf{D_X}(X_1X_2)\cdot d\mathbf{X} + \frac{1}{2}\int_{t_k}^{t_{k+1}}\left(\mathbf{GG}^T\right):\mathbf{D_X^2}(X_1X_2)dt.\right) \\
&\quad = F_1F_2T^2 + (F_1G_2^d + F_2G_1^d)TW^d(T) + G_2^pG_1^s(W^p(T)W^s(T)) \\
&\quad = F_1T\left(F_2T + G_2^dW^d(T)\right) + G_1^pW^p(T)\left(F_2T + G_2^dW^p(T)\right) \\
&\quad = \left(F_1T + G_1^pW^p(T)\right)\left(F_2T + G_2^dW^d(T)\right) \\
&\quad = X_1(T)X_2(T).
\end{aligned}$$

This equality proves the product rule for stochastic processes. □

The product rule for stochastic processes is also *almost* the product rule for deterministic processes. There is just a correction term, like I have shown before in Equation (2.17).

I would also like to note that if we blindly did a Taylor's Approximation on this differential term, we would get the following:

$$\begin{aligned} d(X_1X_2) &= \nabla_X(X_1X_2)\cdot d\mathbf{X} + \frac{1}{2}\left(d\mathbf{X}\right)^T \nabla_X^2(X_1X_2)d\mathbf{X} + ... \\ &= \nabla_X(X_1X_2)\cdot d\mathbf{X} + dX_1dX_2 + ... \\ &= \nabla_X(X_1X_2)\cdot d\mathbf{X} + \left(F_1dt + G_1^i dW^i\right)\left(F_2dt + G_2^j dW^j\right) + ... \\ &= \nabla_X(X_1X_2)\cdot d\mathbf{X} + F_1F_2(dt)^2 + F_1G_2^i dtdW_i + F_2G_1^i dtdW^i + G_1^iG_2^j dW^i dW^j + ... \end{aligned}$$

Comparing this with what we just proved to be the actual stochastic product rule, we see that in order for the two things to be equal, we must have the following:

$$\begin{cases} dW^i dW^j = \delta_{ij}dt \\ dW^i dt = 0 \\ dtdt = 0. \end{cases} \tag{2.24}$$

The reason that this is important to note, is because these are often given as properties of Wiener Processes. Some of the intuition for the first equality actually comes from the quadratic variation of Wiener Processes. These, of course, are not actual equalities, but more of *heuristic* equalities.

These heuristics also give a somewhat easier way of remembering the product rule: Simply apply Taylor's theorem and then use the above rules for simplifying the dX_1dX_2 term.

At this point, I have introduced everything necessary for proving Itô's Lemma. The product rule was the last necessary piece to be able to prove this lemma, so with this in hand, it is time to move on to the most important stochastic lemma in this thesis.

2.3.4 Itô's Lemma

Itô's Lemma is the core lemma that will be necessary in deriving the Black-Scholes Model. The reason is because the function that I will be looking for is a function of both a stochastic process

and of time. Therefore, to find the differential of this function, I will need some sort of "stochastic chain rule". This is, in fact, exactly what Itô's Lemma is: a stochastic chain rule.

Lemma 10 (Itô's Lemma)**.** *Suppose* $\left\{W^k\right\}_{1\leq k\leq m}$ *are i.i.d Wiener Processes. Suppose also that*

$$dX_i = F_i dt + G_i^k dW^k \qquad i = 1, 2, ..., n$$

are n *stochastic processes where* $F_i \in L^1(0,T)$ *and* $G_i^k \in L^2(0,T)$ *for* $i \in \{1,2\}$ *and* $k \in \{1, ..., m\}$ *as in Lemma 9.*

Now, suppose that $u(t, \mathbf{X}(\mathbf{t}))$ *is a function of both* $\mathbf{X}(\mathbf{t}) = (X_1(t), ..., X_n(t))^T$ *and* t*. Then,*

$$d(u(t, \mathbf{X}(\mathbf{t})) = u_t dt + \nabla_X(u(t, \mathbf{X}(\mathbf{t})) \cdot d\mathbf{X} + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(u(t, \mathbf{X}(\mathbf{t}))dt. \tag{2.25}$$

Proof. This lemma has been shown in the [15] for the one dimensional case, but I have extended this to the multi-dimensional case.

I will assume that $u(t, \mathbf{X}(t))$ is a polynomial in $\mathbf{X}$ and t. Then, the proof can be extended to apply to any analytic function u through approximations. The approximation step will not be shown in this thesis because it is beyond the scope of this thesis.

First, suppose that v is only a monomial in $\mathbf{X}$. Then,

$$v(\mathbf{X}) = X_1^{k_1} \cdot X_2^{k_2} \cdot ... \cdot X_n^{k_n}.$$

I claim that the differential of v is

$$dv = \nabla_X(v) \cdot d\mathbf{X} + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(v)dt. \tag{2.26}$$

I will show this through induction on $k_1, k_2, ..., k_n$, however it is enough to simply show this through induction on k_1 because the other k_i's are similar.

The base case here is trivial. If $k_1, ..., k_n = 0$, then $v(\mathbf{X}) = 1$ and

$$dv = \nabla_X(1) \cdot d\mathbf{X} + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(1)dt = 0.$$

For the more interesting case, suppose that

$$\begin{cases} u(\mathbf{X}) = X_1^{k_1} \cdot X_2^{k_2} \cdot ... \cdot X_n^{k_n}, \quad \text{and} \\ d(u) = \nabla_X(u) \cdot d\mathbf{X} + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(u)dt. \end{cases}$$

Consider $v(\mathbf{X}) = X_1 \cdot u$. Then, because I would like to use the product rule, I will let $Y^1 = X_1$ and $Y^2 = u$.

At this point, because I will need the coefficients in front of each dW^k for $k \in \{1, ..., m\}$, then it will be helpful to write both dY^1 and dY^2 slightly differently. I will start with dY^2 first, because it will give insight into how to rewrite dY^1 in a helpful way. So, I have

$$\begin{aligned} dY^2 &= \nabla_X(u) \cdot d\mathbf{X} + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(u)dt \\ &= \frac{\partial}{\partial X_i}\left(Y^2\right) \cdot \left(F_i dt + G_i^k dW^k\right) + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(u)dt \\ &= \left[\frac{\partial}{\partial X_i}\left(Y^2\right) F_i + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(u)\right] dt + \frac{\partial}{\partial X_i}\left(Y^2\right) G_i^k dW^k. \end{aligned}$$

The important term that I am looking for is the coefficient in front of each dW^k, which in this case is $\frac{\partial}{\partial X_i}\left(Y^2\right) G_i^k$.

The term dY^1 is much easier, but I wanted to look at dY^2 first because it will allow me to write the coefficient in front of each dW^k term in a consistent way. So, for dY^1 I have

$$\begin{aligned} dY^1 &= F_1 dt + G_1^k dW^k \\ &= F_1 dt + \frac{\partial}{\partial X_i}\left(X_1\right) G_i^k dW^k \\ &= F_1 dt + \frac{\partial}{\partial X_i}\left(Y^1\right) G_i^k dW^k. \end{aligned}$$

Now, putting all of this together, I see that the coefficients in front of each dW^k is

$$\begin{cases} \frac{\partial}{\partial X_i}\left(Y^1\right) G_i^k & \text{for } dY^1 \text{ and} \\ \frac{\partial}{\partial X_i}\left(Y^2\right) G_i^k & \text{for } dY^2. \end{cases}$$

Therefore, I will construct a matrix for the random variables Y^1 and Y^2 that is very similar to the matrix $\mathbf{G}$ in the product rule. This matrix will contain the coefficients in front of each dW^k term, which is why I went through all of the above work. I will call this matrix $\mathbf{K}$, and it is given by

$$\mathbf{K}_{ij} = \frac{\partial}{\partial X_k}\left(Y^i\right) G_k^j. \tag{2.27}$$

This matrix was important to construct because it will allow me to use Lemma 9 and apply the product rule to find the differential $dv = d\left(Y^1 Y^2\right)$. So, using the product rule, I have

$$d\left(v\right) = \nabla_Y(v) \cdot d\mathbf{Y} + \frac{1}{2}\left(\mathbf{K}\mathbf{K}^T\right) : \nabla_Y^2(v) dt. \tag{2.28}$$

The simplification of this will follow much of the logic used in Lemma 9. I will expand Equation (2.28) to get:

$$\begin{aligned} d(v) &= Y^2 dY^1 + Y^1 dY^2 + \frac{1}{2}\mathbf{K}_{ik}\mathbf{K}_{jk}\left(1 - \delta_{ij}\right) dt \\ &= u dX_1 + X_1\left(\nabla_X(u) \cdot d\mathbf{X} + \frac{1}{2}(\mathbf{G}\mathbf{G}^T) : \nabla_X^2(u) dt\right) + \frac{1}{2}\mathbf{K}_{ik}\mathbf{K}_{jk}\left(1 - \delta_{ij}\right) dt. \end{aligned} \tag{2.29}$$

Note that

$$\begin{aligned} \nabla_X(v) \cdot d\mathbf{X} &= \nabla_X\left(X_1 \cdot u\right) \cdot d\mathbf{X} \\ &= \left(X_1 \cdot \nabla_X(u) + u \cdot \nabla_X\left(X_1\right)\right) \cdot d\mathbf{X} \\ &= X_1\left(\nabla_X(u) \cdot d\mathbf{X}\right) + u dX_1. \end{aligned} \tag{2.30}$$

Notice that this expression appears in Equation (2.29). So, substituting (2.30) into Equation (2.29), I get

$$d(v) = \nabla_X(v) \cdot d\mathbf{X} + \frac{1}{2} X_1 \left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(u) dt + \frac{1}{2} \mathbf{K}_{ik}\mathbf{K}_{jk} \left(1 - \delta_{ij}\right) dt.$$

Here, I am very close to the result that I am looking for. I only need to show that the last two terms combine properly. So, looking at the last two terms, I see

$$\begin{aligned} &\frac{1}{2} X_1 \left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(u) dt + \frac{1}{2} \mathbf{K}_{ik}\mathbf{K}_{jk} \left(1 - \delta_{ij}\right) dt \\ &= \frac{1}{2} Y^1 G_i^k G_j^k \frac{\partial^2}{\partial X_i \partial X_j} \left(Y^2\right) dt + \frac{1}{2} \frac{\partial}{\partial X_l} \left(Y^i\right) \frac{\partial}{\partial X_r} \left(Y^j\right) G_l^k G_r^k \left(1 - \delta_{ij}\right) dt. \end{aligned} \tag{2.31}$$

Here, it is important to note that $i, j \in \{1, 2\}$ because there are only Y^1 and Y^2. Therefore, I will explicitly write out the sum over i and j. In order to make the notation a bit simpler, I will also adopt the notational choice to let $Y_{X_i} = \frac{\partial}{\partial X_i}(Y)$. So, doing this explicit summing, I have

$$\begin{aligned} &\frac{1}{2} Y^1 G_i^k G_j^k Y^2_{X_i X_j} dt + \frac{1}{2} Y^1_{X_l} Y^2_{X_r} G_l^k G_r^k dt + \frac{1}{2} Y^2_{X_l} Y^1_{X_r} G_l^k G_r^k dt \\ &= \frac{1}{2} G_l^k G_r^k \left(Y^1 Y^2_{X_l X_r} + Y^1_{X_l} Y^2_{X_r} + Y^1_{X_r} Y^2_{X_l}\right) dt. \end{aligned} \tag{2.32}$$

This expression is starting to look much simpler, but before I can do the final simplification, I must think about what I am really looking for. I am trying to show that this expression above equals $\frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(v) dt$. It may be helpful then, to expand this and see what it looks like:

$$\begin{aligned} \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(v) dt &= \frac{1}{2} G_i^k G_j^k \frac{\partial^2}{\partial X_i \partial X_j} \left(Y^1 Y^2\right) dt \\ &= \frac{1}{2} G_i^k G_j^k \frac{\partial}{\partial X_i} \left(Y^1_{X_j} Y^2 + Y^2_{X_j} Y^1\right) dt \\ &= \frac{1}{2} G_i^k G_j^k \left(Y^1_{X_i X_j} Y^2 + Y^1_{X_j} Y^2_{X_i} + Y^2_{X_j} Y^1_{X_i} + Y^2_{X_i X_j} Y^1\right) \\ &= \frac{1}{2} G_i^k G_j^k \left(Y^1_{X_j} Y^2_{X_i} + Y^2_{X_i} Y^1 X_j + Y^2_{X_i X_j} Y^1\right) dt, \end{aligned}$$

where the last line comes from the fact that $Y^1 = X_1$, and thus any second partial derivative will be zero. Notice that this is exactly what I have in Equation (2.32). So, after substituting what I have just found into Equation (2.31), I get

$$d(v) = \nabla_X(v) \cdot d\mathbf{X} + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(v)dt.$$

Therefore, I have shown the differential of a polynomial in $\mathbf{X}$. This was quite a bit of work and a long detour, but I am nearly done with the proof of Itô's Lemma. Recall that this lemma is concerned with a function of both $\mathbf{X}$ and of t. I therefore need to build upon the polynomial in $\mathbf{X}$ that I have created above, and include the time variable. So, now suppose that

$$v\left(\mathbf{X}, t\right) = f\left(\mathbf{X}\right) g\left(t\right),$$

where f and g are both polynomials. Because g is a deterministic function, then there is no Itô correction term. Therefore, the differential of this function is the "usual" product rule. So, I have the differential of v to be

$$\begin{aligned} d(v) &= d\left(f\left(\mathbf{X}\right)\right) g(t) + d\left(g(t)\right) f\left(\mathbf{X}\right) \\ &= g(t)\left(\nabla_X(f) \cdot d\mathbf{X} + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(f)dt\right) + g'(t) f\left(\mathbf{X}\right) dt \\ &= v_t dt + \nabla_X(v) \cdot d\mathbf{X} + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_X^2(v)dt. \end{aligned}$$

So, I have shown that the result from Itô's Lemma is true for polynomials. The last step, which I have mentioned is beyond the scope of this thesis, is to rigorously show that this result is true for general analytic functions. The idea is that any analytic function can be approximated by polynomials, so this result is in fact true for any general function. This then completes the proof of Itô's Lemma. □

2.4 Chapter Summary

In this chapter, I introduced the basic building block of stochastic processes: *the Wiener Process*. This process was crucial to many of the theory introduced in this chapter.

I then talked about some of the basic properties of Wiener Processes, with one of the most important properties being the *Quadratic Variation* of Wiener Processes. This property was necessary to being able to define the next important thing in this chapter: the *stochastic integral*.

I defined the stochastic integral through a Riemann-like sum in a similar fashion to deterministic integrals. This sum was able to converge because of the quadratic variation property of the Wiener Process, however the sum converged to different values based on the evaluation point. This led me to have to choose an evaluation point to define the integral, which in my thesis I chose to be the left endpoint of each interval. This definition turned out to be what is called the *Itô Integral*.

With the stochastic integral defined, I defined the *stochastic differential* even though Wiener Processes are nowhere-differentiable. I then gave several lemmas that ultimately led to proving the stochastic product rule and chain rule. The latter of these is what is called *Itô's Lemma*, which is the most important stochastic lemma necessary to derive the Black-Scholes Model.

With all of this stochastic background understood, I can now move on to the Black-Scholes Model.

Chapter 3

Black-Scholes Model

As I mentioned in Chapter 1, the Black-Scholes Model is a mathematical model used to accurately price options. This is important because, without an accurate way of pricing the options, then there would be arbitrage opportunities in the market.

In reality, there are actually many different flavors of options, and they all behave slightly differently. For this thesis, I will be focused on European Call Options. These give the owner the *right* to purchase an asset at the expiration time, but not at any time before. From this point on, when I use the term "option", I am implicitly referring to the European Call Option.

This chapter will focus on the derivation of the Black-Scholes Model. This thesis is mainly focused on what is called the **Basket Case Model**, which is mainly a generalization of the model that was proposed in [8]. This Basket Case Model is about pricing options that are for more than one underlying asset. This can be thought of as the *multi-dimensional* version of the Black-Scholes Model.

I will start the chapter by giving a derivation of the original Black-Scholes Model as proposed in [8]. I will sometimes refer to this as the *one-dimensional* version of the Black-Scholes Model. Most of the derivation logic is taken from this paper, but it is helpful to walk through it in order to gain a better understanding of the more complicated Basket Case Model. The reason is because the one-dimensional derivation will motivate some of the derivation that I will do in the multi-dimensional version.

After deriving the One-Dimensional Black-Scholes Model, I will generalize this process to higher dimensions, or more underlying assets. This will involve some linear algebra techniques in order to use the theorems discussed in Chapter 2.

This chapter will end then, with a formal statement of the Basket-Case Black-Scholes Model that will be solved in later chapters.

3.1 One-Dimensional Black-Scholes Model

This section will discuss the derivation of the One-Dimensional Black-Scholes Model, as well as discuss the model itself. As mentioned in the introduction of this chapter, it is important to discuss and understand the one-dimensional model before I can discuss the Basket Case Model. The reason is that the derivation in both cases is very similar, but the process is slightly easier to understand initially in the one-dimensional case.

In this section I will also discuss the boundary conditions for this model. These are often just as interesting and important as the PDE itself is, so it is important that I talk about them as well.

At the end of this section, I will have a formal statement for the one-dimensional Black-Scholes Model.

3.1.1 Delta Hedging

The derivation of the one-dimensional model will depend heavily on the idea of *delta hedging*. This is the idea behind the original derivation by Fischer Black and Myron Scholes [8].

From the financial perspective, delta hedging can be understood as trying to remove the risk in the price fluctuations of the option by also investing in some amount of the underlying asset to cancel out any price fluctuations of the option induced by changes in price of the underlying asset [12].

If the risk due to the price fluctuations of the underlying asset are removed, then the portfolio should be expected to grow at the risk free interest rate. Then, it becomes much easier to value the portfolio, and therefore the option.

The main idea with this approach is that a portfolio consisting of the option and some amount of the underlying asset will be constructed in such a way as to remove risk. Then, the portfolio can be valued fairly easily because it is expected to grow at the risk free interest rate. Finally, with the portfolio valued, the value of the option can be deduced.

3.1.2 One-Dimensional Derivation

With an understanding of how delta hedging works, I can now derive the one-dimensional Black-Scholes model. Because this thesis is a mathematics thesis, however, I will take a more mathematical approach to delta hedging.

First, I need to state an assumption on the underlying asset. For this model, I will be assuming that the value of the underlying asset, given by S, follows a Geometric Brownian Motion. This was not explicitly referenced by this name in Chapter 2, but was used in many of the lemmas. Formally, this means that

$$dS = rSdt + \sigma SdW, \tag{3.1}$$

where

- r is the risk-free interest rate and
- σ is the volatility of the asset.

With that assumption made, I can continue with the derivation of the value of the option. I will let $V(S,t)$ represent the value of the option. What I will do, is approximate this function at the current stock price, S_k, and current time, t_k. Using a first order Taylor Expansion, I have

$$V(S,t) \approx V(S_k,t_k) + V_S(S_k,t_k)(S-S_k) + V_t(S_k,t_k)(t-t_k). \tag{3.2}$$

The function $V(S,t)$ has some amount of randomness that is induced by the randomness of the value of the underlying asset, given by S. If I could remove the randomness here, then the value would be deterministic. So, what I will do is create another function, $P(S,t)$, that represents the value of a portfolio with no randomness. All I have to do to create this function is take $V(S,t)$ and subtract the random components. So, using my approximation in Equation (3.2), I have

$$P(S,t) = V(S,t) - V_S(S_k,t_k)S. \tag{3.3}$$

Notice that this doesn't *fully* remove the randomness, because I am only subtracting the first order term associated with the randomness. As long as the price doesn't fluctuate *too* much, however, this approximation will do.

Since I now have an equation for the value of a portfolio without randomness, I should expect this portfolio to grow at the risk free interest rate. Mathematically, this means

$$dP = rPdt. \tag{3.4}$$

In the above equation, I have suppressed the arguments to the function for clarity. From this point onward, I will continue to suppress the arguments of the functions.

I can also express this differential in another way, based on the definition of P as

$$dP = dV - V_S dS. \tag{3.5}$$

Setting Equations (3.4) and (3.5) equal, I have

$$\begin{aligned} rPdt &= dV - V_S dS \\ \implies 0 &= dV - V_S dS + rSV_S dt - rV dt, \end{aligned} \tag{3.6}$$

where I substituted $P = V - V_S S$ and moved all terms to the right hand side.

At this point, I can finally use the results from Chapter 2. In particular, I will use Itô's Lemma to rewrite the dV term. Recall, that I can rewrite this as

$$dV = V_t dt + V_S dS + \frac{1}{2}\sigma^2 S^2 V_{SS} dt. \tag{3.7}$$

Substituting this into Equation (3.6), I get:

$$\begin{aligned} 0 &= V_t dt + V_S dS + \frac{1}{2}\sigma^2 S^2 V_{SS} dt - V_S dS + rSV_S dt - rV dt \\ &= (V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV)dt. \end{aligned} \tag{3.8}$$

The last step here is to realize that in order for Equation (3.8) to equal 0, the following must be true

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0. \tag{3.9}$$

At this point, I have arrived at the one-dimensional Black-Scholes PDE.

The key with this argument was to use delta hedging, which is really just a first order approximation to remove the randomness, in order to deduce the value of the option. Of course, I haven't yet found the value of the option yet, but I do have a PDE with *no random components* that describes the value of the option. This is a purely deterministic PDE that I can solve to get the value of the option.

What I will do next, is describe the boundary conditions for this model.

3.1.3 One-Dimensional Boundary and Initial Conditions

To describe the one-dimensional model fully, I need to define the boundary conditions as well. These are just as important and interesting as the PDE itself, so I will go into the detail of them in this section.

For the first boundary condition, I will consider what the value of the option should be when the underlying asset is worthless. In this case, the holder of the option will never want to exercise the option. The reason is because the strike price is non-zero, and no reasonable investor would want to pay the strike price for a worthless asset. Therefore, the reasonable investor would never exercise the option in this case, and it is providing no value to the investor. This leads to the first

boundary condition:

$$V(0,t) = 0, \qquad \forall t \in [0,T],$$

where T is the expiration time.

The next boundary condition should describe the solution at the other extreme value for $S(t)$. Because the value of the underlying asset is limitless (in theory), then we must describe what the solution does as $S \to \infty$. In this case, the corresponding put option will almost certainly not be exercised. Therefore, by the Put-Call Parity, I can deduce the value of the call option to be:

$$V(S,t) \to S - Ke^{-rt}, \qquad S \to \infty, \forall t \in [0,T],$$

where T is again the expiration time.

The last condition then, is the final condition. I am not using an initial condition here, because the value of the option at expiration time is clearer to describe. The value of the option at expiration time is simply the amount of payoff that we will receive. If the value of the underlying asset is above the strike price K, then the holder will exercise the option and the payoff will simply be the difference between the current price and the amount paid for the asset (the strike price). If the value of the asset is below the strike price, then the holder will not exercise the option, and will simply purchase the asset at the current market price, thus making the option worthless. This leads to the final condition of:

$$V(S,T) = \max(S - K, 0), \qquad \forall S \in \Omega,$$

where T is the expiration time, K is the strike price, and Ω is the set of possible values that the asset can have.

With the boundary conditions defined, as well as the PDE, I can finally move on to formally stating the one-dimensional Black-Scholes model.

3.1.4 One-Dimensional Model Statement

With an understanding of the PDE itself, as well as the boundary conditions, I can actually state the model.

Model 1 (One-Dimensional Black-Scholes Model)**.** *Suppose that a call option, with strike price K and expiration time T, consists of one underlying asset whose value follows a Geometric Brownian Motion given by*

$$dS = rSdt + \sigma SdW,$$

where r represents the risk-free interest rate and σ represents the volatility of the asset's value. Also suppose that Ω is the set of possible values of the underlying asset. Then the value for the call option, $V(S,t)$, is modeled by

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0, \qquad \forall S \in \Omega, t \in [0,T], \tag{3.10}$$

with boundary conditions given by

$$V(0,t) = 0, \qquad \forall t \in [0,T] \tag{3.11}$$

$$V(S,t) \to S - Ke^{-rt} \qquad S \to \infty, \forall t \in [0,T] \tag{3.12}$$

$$V(S,T) = \max(S-K,0) \qquad \forall S \in \Omega. \tag{3.13}$$

Although my thesis will not be focusing on solving this version of the Black-Scholes model, it is useful to state it here. This is because the Basket-Case version will use many of the ideas from this model.

3.2 Basket-Case Black-Scholes Model

This section will focus on the main problem that this thesis is concerned with solving. With an understanding of both stochastic calculus and call options, I can start to discuss the more interesting Basket-Case Black-Scholes Model. As a reminder, a Basket-Case option is an option that consists

of more than one underlying asset. This will cause the value of the option to now depend on the values of not one, but several assets, all of which will have some randomness to their value. So, this Basket-Case model will be a model that prices one of these Basket-Case options.

These have a very good use in today's markets, just as the one-dimensional (or single asset) models do. Consider a multinational corporation that needs to do business in various currencies. This corporation will then need to manage the risk of doing business in several different currencies. One thing they could do to protect themselves, would be to use a Basket-Case option that has the various currencies as the underlying assets. This way, instead of purchasing options for each different currency, the corporation can simply purchase a single option that covers all of the currencies. Some of the benefits of this approach are that only one option needs to be managed instead of many, and there are fewer transaction costs.

This section will focus on the derivation and statement of the Basket-Case Black-Scholes Model. I will start by rewriting the models for each of the underlying assets in order to use Itô's Lemma. Then, I will propose a vector model for the underlying assets. With a vector model, I will be able to derive the Basket-Case model. Finally, I will state the Basket-Case Black-Scholes Model.

3.2.1 Asset Modification

One of the assumptions of the Basket-Case model, just like in the one-dimensional model, is that each underlying asset follows a Geometric Brownian Motion. This means that, given a Basket-Case option with N underlying assets, I will model each of the underlying asset prices by

$$dS_i = rS_i dt + \sigma_i S_i dW_i, \qquad \forall i \in (1, 2, 3, ..., N), \tag{3.14}$$

where σ_i is the volatility of the ith underlying asset, dW_i is the Brownian Motion that will contribute to the randomness in the change in value of the asset and N is the number of underlying assets in the option.

Based on the derivation of the one-dimensional model in Section 3.1, I may be tempted to construct a risk-less portfolio and then use Itô's Lemma to rewrite the differential of the value of the option. Unfortunately, this will not work. The reason is because the model for each of the underlying assets given in Equation (3.14) does not match the assumptions in order to use Itô's Lemma. More specifically, each of the dW_i's is not necessarily independent from the rest. Physically, this is because the "randomness" in the changes in price for each underlying asset is not necessarily independent. A change in the price of one asset very well might influence a change in price of another asset.

In order to use Itô's Lemma then, I will need to adjust the models for each of the underlying assets to make this "randomness" independent. To do this, I will take some inspiration from linear algebra and do something similar to building orthonormal basis vectors from a spanning set.

In order to correct my model in Equation (3.14), I will need to impose the following conditions on the dS_i terms:

1. Each dW_i needs to be represented as

$$dW_i = \sum_{j=0}^{N} \alpha_{ij} d\widehat{W}^j,$$

   where $\left\{ d\widehat{W}^j \right\}_{j=1,\ldots,N}$ are differentials of a Wiener Process and are pairwise independent and each α_{ij} is simply a real number, and

2. for each dW_i, I should still have

$$dW_i \sim \mathcal{N}(0, dt)$$

   and lastly

3. the correlation between dS_i and dS_j should equal a prescribed value, given by ρ_{ij}.

These restrictions are ultimately a problem of finding some very special coefficients that will allow me to represent each dW_i term as a linear combination of linear independent Wiener Process

differentials in a convenient way. With this new formulation of each dW_i term, and therefore each dS_i term, I will be able to use Itô's Lemma. This is because the Wiener Process differentials in the model will be pairwise independent.

The algorithm I am going to use to solve for these coefficients is a modified version of the Gram–Schmidt process.

First, I will start with a set of pairwise independent Wiener Process differentials

$$\left\{d\widehat{W}^j\right\}_{j=1,\ldots,N}. \tag{3.15}$$

These will be a starting point for rewriting the models for the asset differentials.

Next, I will simply set the first asset differential to

$$dS_1 = rS_1dt + \sigma_1 S_1 d\widehat{W}^1. \tag{3.16}$$

To get the second asset differential, I will first impose the first condition to rewrite it as

$$dS_2 = rS_2dt + \sigma_2 S_2\left(\alpha_{21}d\widehat{W}^1 + \alpha_{22}d\widehat{W}^2\right). \tag{3.17}$$

Imposing the second condition, I see that

$$\alpha_{21}d\widehat{W}^1 + \alpha_{22}d\widehat{W}^2 \sim \mathcal{N}\left(0, dt\right).$$

Because of the linearity of the mean, and because each $d\widehat{W}^i \sim \mathcal{N}(0, dt)$, then the above sum is already normally distributed with mean zero. This simply leaves a restriction on the variance. So,

$$\begin{aligned}
\text{var}\left(\alpha_{21}d\widehat{W}^1 + \alpha_{22}d\widehat{W}^2\right) &= \text{var}\left(\alpha_{21}d\widehat{W}^1\right) + \text{var}\left(\alpha_{22}d\widehat{W}^2\right) \\
&= \alpha_{21}^2\,\text{var}\left(d\widehat{W}^1\right) + \alpha_{22}^2\,\text{var}\left(d\widehat{W}^2\right) \\
&= \left(\alpha_{21}^2 + \alpha_{22}^2\right)dt \\
\implies \alpha_{21}^2 + \alpha_{22}^2 &= 1.
\end{aligned} \tag{3.18}$$

Finally, imposing the third condition I see that

$$\begin{aligned}
\rho_{12} = \text{Corr}\left(dS_1, dS_2\right) &= \frac{\text{Cov}\left(dS_1, dS_2\right)}{\sqrt{\text{var}\left(dS_1\right)\text{var}\left(dS_2\right)}} \\
&= \frac{\text{Cov}\left(d\widehat{W}^1, \alpha_{21}d\widehat{W}^1 + \alpha_{22}d\widehat{W}^2\right)}{\sqrt{\text{var}\left(d\widehat{W}^1\right)\text{var}\left(\alpha_{21}d\widehat{W}^1 + \alpha_{22}d\widehat{W}^2\right)}} \\
&= \frac{\text{Cov}\left(d\widehat{W}^1, \alpha_{21}d\widehat{W}^1\right) + \text{Cov}\left(d\widehat{W}^1, \alpha_{22}d\widehat{W}^2\right)}{\sqrt{dt\cdot\left(\alpha_{21}+\alpha_{22}\right)dt}} \\
&= \frac{\alpha_{21}\,\text{var}\left(d\widehat{W}^1\right)}{dt} \\
&= \alpha_{21}.
\end{aligned} \tag{3.19}$$

Thus, I see that $\alpha_{21} = \rho_{12}$. To find α_{22}, I simply need to substitute this result into Equation (3.18)

$$\begin{aligned}
\alpha_{22}^2 &= 1 - \alpha_{21}^2 \\
\implies \alpha_{22} &= \sqrt{1-\rho_{12}^2}.
\end{aligned} \tag{3.20}$$

Putting both Equations (3.23) and (3.20) together, I can substitute these results into Equation (3.17) to get the full model for the second asset differential

$$dS_2 = rS_2dt + \sigma_2 S_2 \left(\rho_{12} d\widehat{W}^1 + \sqrt{1-\rho_{12}^2} d\widehat{W}^2 \right). \tag{3.21}$$

Because my thesis focuses on solving the three-dimensional Basket Case Model, I will continue one more iteration of this process to get the model for the third asset differential.

So, using the first condition I have

$$dS_3 = rS_3dt + \sigma_3 S_3 \left(\alpha_{31} d\widehat{W}^1 + \alpha_{32} d\widehat{W}^2 + \alpha_{33} d\widehat{W}^3 \right). \tag{3.22}$$

Imposing the second condition will lead me again to the restriction that the variance of dS_3 should equal dt. Thus,

$$\begin{aligned} \text{var}\left(\alpha_{31} d\widehat{W}^1 + \alpha_{32} d\widehat{W}^2 + \alpha_{33} d\widehat{W}^3 \right) &= \left(\alpha_{31}^2 + \alpha_{32}^2 + \alpha_{33}^2 \right) dt \\ \implies \alpha_{31}^2 + \alpha_{32}^2 + \alpha_{33}^2 &= 1. \end{aligned} \tag{3.23}$$

Finally, in order to impose the third condition, I need to consider two correlation coefficients: ρ_{13} and ρ_{23}. So, I will consider ρ_{13} first:

$$\begin{aligned} \rho_{13} &= \frac{\text{Corr}\left(dS_1, dS_3\right)}{\sqrt{\text{var}\left(dS_1\right) \text{var}\left(dS_3\right)}} \\ &= \frac{\text{Corr}\left(d\widehat{W}^1, \alpha_{31} d\widehat{W}^1 + \alpha_{32} d\widehat{W}^2 + \alpha_{33} d\widehat{W}^3 \right)}{\sqrt{dt \cdot dt}} \\ &= \frac{\alpha_{31} dt}{dt} \\ &= \alpha_{31}. \end{aligned} \tag{3.24}$$

Similarly, for ρ_{23}, I have

$$\begin{aligned}
\rho_{23} &= \frac{\operatorname{Corr}\left(dS_2, dS_3\right)}{\sqrt{\operatorname{var}\left(dS_2\right)\operatorname{var}\left(dS_3\right)}} \\
&= \frac{\operatorname{Corr}\left(\rho_{12}d\widehat{W}^1 + \sqrt{1-\rho_{12}^2}d\widehat{W}^2, \alpha_{31}d\widehat{W}^1 + \alpha_{32}d\widehat{W}^2 + \alpha_{33}d\widehat{W}^3\right)}{dt} \\
&= \frac{\left(\rho_{12}\alpha_{31} + \sqrt{1-\rho_{12}^2}\alpha_{32}\right)dt}{dt} \\
&= \rho_{12}\alpha_{31} + \sqrt{1-\rho_{12}^2}\alpha_{32}
\end{aligned} \tag{3.25}$$

Now, I will substitute the result in Equation (3.24) into the above equation to obtain

$$\alpha_{32} = \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}. \tag{3.26}$$

To get the last coefficient, I simply need to substitute the results from Equations (3.26) and (3.24) into Equation (3.23)

$$\alpha_{33} = \sqrt{1 - \rho_{13}^2 - \left(\frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}\right)^2}. \tag{3.27}$$

Finally, substituting the results from Equations (3.23), (3.24) and (3.26) into the model in Equation (3.22)

$$\begin{aligned}
dS_3 &= rS_3dt \\
&+ \sigma_3 S_3\left(\rho_{13}d\widehat{W}^1 + \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}d\widehat{W}^2 + \sqrt{1 - \rho_{13}^2 - \left(\frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}}\right)^2}d\widehat{W}^3\right).
\end{aligned} \tag{3.28}$$

This process will continue until a model for all N underlying assets is constructed. For my three-asset case, I have the following models for the underlying assets:

$$\begin{aligned}
dS_1 &= rS_1dt + \sigma_1 S_1 d\widehat{W}^1 \\
dS_2 &= rS_2dt + \sigma_2 S_2 \left(\rho_{12} d\widehat{W}^1 + \sqrt{1-\rho_{12}^2}\, d\widehat{W}^2 \right) \\
dS_3 &= rS_3dt \\
&\quad + \sigma_3 S_3 \left(\rho_{13} d\widehat{W}^1 + \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} d\widehat{W}^2 + \sqrt{1-\rho_{13}^2 - \left(\frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} \right)^2}\, d\widehat{W}^3 \right).
\end{aligned} \tag{3.29}$$

The benefit of rewriting my models for each underlying asset is that the Wiener Process differentials are now independent, and the correlation between any two asset differentials, dS_i and dS_j, is ρ_{ij}. In this way, I have prescribed the pairwise correlation between each asset differential, and I have satisfied the conditions to be able to use Itô's Lemma.

Before deriving the Basket Case Black-Scholes Model, I will do one more thing to my asset models to make the derivation slightly easier: I will combine the three asset models above into one vector model. This combining will make the use of the multi-dimensional Itô's Lemma in Equation (2.25) easier.

3.2.2 Multi-Asset Vector Model

The previous section described how to reformulate the underlying asset differential models into a way that will allow me to use Itô's Lemma. This section will focus on combining the models in Equation (3.29) into one single vector model. The reason for doing this is to allow the easy use of Equation (2.25).

Recall that a convenient way to represent the differentials of several stochastic processes is to use what I will call the *vector model* of the stochastic differentials. This vector form is given by

$$d\mathbf{X} = \mathbf{R}\left(\mathbf{X}, t\right) dt + \mathbf{G}\left(\mathbf{X}, t\right) d\mathbf{W}(t), \tag{3.30}$$

where

$$\begin{cases} \mathbf{R}_i = F_i & i = 1, \ldots, N \\ \mathbf{G}_{ij} = G_i^j & i, j = 1, \ldots, N \\ d\mathbf{W}_i = d\widehat{W}^i & i = 1, \ldots, N. \end{cases}$$

Here, F_i and G_i^j are the same as in Lemma 9 and N is the number of underlying assets that are being modeled. Recall that F is a vector of random variables, and G is a matrix with random variables as entries.

With this definition in mind, I will create a vector model for the models constructed in Equation (3.29) as follows

$$d\mathbf{S} = \mathbf{R}dt + \mathbf{G}d\mathbf{W}. \tag{3.31}$$

Here I have suppressed the dependencies on $\mathbf{S}$ and t for clarity.

This model is not quite complete however, because I have not defined what $\mathbf{R}$ and $\mathbf{G}$ are. These are not too difficult to get, luckily, as $\mathbf{R}$ simply describes the coefficients in front of the dt term and $\mathbf{G}$ describes the coefficients in front of the dW_i terms.

The $\mathbf{R}$ term is the easiest to extract, as this is simply given by

$$\mathbf{R} = \begin{bmatrix} rS_1 \\ rS_2 \\ rS_3 \end{bmatrix}. \tag{3.32}$$

The $\mathbf{G}$ term is not quite as pretty, but after looking at the terms in front of each dW^i term in Equation (3.29), I have

$$\mathbf{G} = \begin{bmatrix} \sigma_1 S_1 & 0 & 0 \\ \sigma_2 \rho_{1,2} S_2 & \sigma_2 \sqrt{1 - \rho_{1,2}^2} S_2 & 0 \\ \sigma_3 \rho_{1,3} S_3 & \sigma_3 \frac{\rho_{2,3} - \rho_{1,2}\rho_{1,3}}{\sqrt{1-\rho_{1,2}^2}} S_3 & \sigma_3 \sqrt{1 - \rho_{1,3}^2 - \frac{(\rho_{2,3} - \rho_{1,2}\rho_{1,3})^2}{1-\rho_{1,2}^2}} S_3 \end{bmatrix}. \tag{3.33}$$

With these two definitions, the vector model for the underlying assets differential is complete. I have only shown the vector model for 3 underlying assets, but this of course can be extended to N underlying assets.

This vector model will allow for the easy application of Itô's Lemma, which is a core lemma in deriving the Basket Case Black-Scholes Model, just as it was essential in deriving the single asset model.

In the next section, I will use the vector model to derive the Basket Case model.

3.2.3 Basket-Case Derivation

I would now like to derive the Basket-Case Black-Scholes Model. Now that I have constructed a model for the underlying assets that will be compatible with Itô's Lemma, I will be able to complete the derivation.

This derivation will follow a fairly similar pattern that was used in Section 3.1.2. I did complete the jump from the single asset model derivation to this Basket-Case model's derivation, with most of the work being the vector and matrix representations of the statements in the single asset derivation. This was particularly useful because it allowed for the concise writing of the derivation, it will show clearly the parallels between the Basket-Case derivation and the single asset model, and it will make the weak formulation much easier in the next chapter. Because most of the logic was borrowed from the single asset model, however, this derivation is mostly credited to the book by Evans [15].

First, I will start with a linear approximation of the value of the Basket-Case Option. If I approximate this function at the current stock prices for each of the underlying assets, $\mathbf{S}^k$, and at the current time, t^k, I have

$$V(\mathbf{S}, t) \approx V\left(\mathbf{S}^k, t^k\right) + \nabla_S(V)\left(\mathbf{S}^k, t^k\right) \cdot \left(\mathbf{S} - \mathbf{S}^k\right) + V_t\left(\mathbf{S}^k, t^k\right)\left(t - t^k\right). \tag{3.34}$$

This is useful because, as long as the prices don't fluctuate *too* much, then this is a good approximation for the value of the Basket-Case option. This approximation also clearly separates out the dependencies on both the asset prices and time.

What I will do then, is create a portfolio, whose price is given by $P(\mathbf{S}, t)$, that will remove the "randomness" in the Basket-Case option's price. To do this, just as in the single asset case, I will subtract away the term that depends on $\mathbf{S}$ from the value of the Basket-Case option to get

$$P(\mathbf{S}, t) = V(\mathbf{S}, t) - \nabla_S(V)\left(\mathbf{S}^k, t^k\right) \cdot \mathbf{S}. \tag{3.35}$$

The usefulness of this specific portfolio, is that its *change* in value is very easy to calculate. The reason for this is because this portfolio should have no risk associated with it, since I have removed the only term from the Basket-Case option's value function that contributes to the risk. Now, because the portfolio P has no risk associated with it's value, then it must grow at the *risk-free interest rate*. In other words,

$$dP = rPdt. \tag{3.36}$$

I have suppressed the dependencies on the asset prices and time here for clarity. I will continue to do this to make things more clear.

Equation (3.36) is not the only way to represent the change in value of the portfolio P, however. I can also use the original definition of the value of P given in Equation (3.34) to get

$$dP = dV - \nabla_S(V) \cdot d\mathbf{S}. \tag{3.37}$$

With two equations for dP, I will set them equal to have

$$\begin{aligned} rPdt &= dV - \nabla_S(V) \cdot d\mathbf{S} \\ \implies 0 &= dV - \nabla_S(V) \cdot d\mathbf{S} + r\nabla_S(V) \cdot \mathbf{S}dt - rVdt, \end{aligned} \tag{3.38}$$

where I substituted $P = V - \nabla_S(V) \cdot \mathbf{S}$ and moved all terms to the right hand side.

At this point, I can use Itô's Lemma to rewrite the dV term as

$$dV = V_t dt + \nabla_S(V) \cdot d\mathbf{S} + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) dt. \tag{3.39}$$

Substituting this into Equation (3.38), I have

$$\begin{aligned} 0 &= V_t dt + \nabla_S(V) \cdot d\mathbf{S} + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) dt - \nabla_S(V) \cdot d\mathbf{S} + r\nabla_S(V) \cdot \mathbf{S} dt - rV dt \\ &= \left(V_t + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) + r\nabla_S(V) \cdot \mathbf{S} - rV\right) dt. \end{aligned} \tag{3.40}$$

The last step in the derivation is, just as in the single asset case, the only way for Equation (3.40) to equal zero, the following must be true by the Fundamental Lemma of Calculus of Variations:

$$V_t + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) + r\nabla_S(V) \cdot \mathbf{S} - rV = 0. \tag{3.41}$$

This is the Basket-Case Black-Scholes PDE.

Just as in the single asset case, the crux of this derivation was delta hedging. It is also worth noting that this derivation so closely parallels the single asset derivation because of the work to write my models into a vector form, and that this derivation was only possible because of the reformulation of the models for each underlying asset that was done in Section 3.2.1.

Next, I will discuss a few of the boundary conditions of the Basket-Case Model. These conditions are just as important as the PDE itself, and are particularly interesting. The particular boundary conditions that are of interest first, are the *left-side* boundary conditions. By left-side, I mean boundary conditions that correspond to one of the underlying assets being equal to zero for the lifetime of the option. As I will discuss in the following section, these particular boundary conditions are very interesting because of how they require me to solve the Basket-Case Model.

3.2.4 Left-Side Basket-Case Boundary Conditions

In order for me to fully state the Basket-Case Black-Scholes Model, I need to state the boundary conditions as well. For this problem, the boundary conditions are particularly interesting because they lead to a peculiar way of solving the PDE.

Suppose that I am trying to model a Basket-Case option with N underlying assets. As described in Section 3.2.3, I can model the value of this option with a PDE. Instead of numerically solving this *one* PDE, I will have to numerically solve *several* lower-dimensional versions of the PDE. The reason is because of the boundary conditions: Each asset leads to a boundary condition that is the solution to a lower-dimensional version of the Basket-Case Model. At the end of this section I will detail exactly how many PDEs will need to be solved, but for now I would like to focus on why the boundary conditions are what they are.

For a Basket-Case option with N underlying assets, then I will need $2N$ boundary conditions, with N of these corresponding to one of the assets being zero for all times t. Let these N boundary conditions be given by

$$V\left(S_1, \ldots, S_{i-1}, 0, S_{i+1}, \ldots, S_N\right) = V_i\left(S_1, \ldots, S_{i-1}, S_{i+1}, \ldots, S_N, t\right), \qquad \forall i \in \{1, \ldots, N\}.$$

Here, the i-th equation corresponds to the boundary condition where the i-th component of $\mathbf{S}$ equals zero for all times t. But, if the value of this asset, S_i, is worthless for all time t, then the value of the option is not impacted by this asset whatsoever. This means that the value of the option depends only on the other $N-1$ assets. In other words, the value of the option when the i-th underlying asset is worthless is simply the solution to the Basket-Case Black-Scholes PDE with the underlying assets being the other $N-1$ assets.

The only 'kink' in the above line of logic is at the dimensions of zero and one. A Basket-Case option with one underlying assets and zero underlying assets doesn't really make sense, so I will have to define what is meant here. For the one-dimensional Basket-Case option, I will define this

to be the single asset option, and for the zero-dimensional Basket-Case option, I will define this to be an option that is always worthless.

The definition for the one-dimensional Basket-Case option follows from the fact that a single asset option is simply a special case of the Basket-Case option. The definition for the zero-dimensional Basket-Case option is because this is the boundary condition for the one-dimensional Basket-Case option when the asset is worthless for all times t. Intuitively, however, this is because if there are no underlying assets, then there is really nothing to trade.

These boundary conditions are extremely important and interesting because in order to solve an N-dimensional Basket-Case Model, I must first solve $N-1$ lower-dimensional problems. And of course, to solve each of the $N-1$ problems, I must solve $N-2$ of their lower-dimensional problems for their boundary conditions.

Therefore, to solve an N-dimensional Basket-Case Model, I will have to first solve for the left-side boundary conditions. Following the pattern above, the number problems that I will have to solve is

$$\#\text{ of problems} = \sum_{i=1}^{N} \binom{N}{i} = 2^N - 1.$$

The left-side boundary conditions for the Basket-Case Model are very important to the model itself, and add some interesting complexity to the problem. Not only does the problem become increasingly more difficult to solve numerically with each additional underlying asset because of the "curse of dimensionality", but there are several nested lower-dimensional Basket-Case problems to solve as well.

With the left-side boundary conditions understood, I will move on to stating the other boundary conditions as well as the initial condition.

3.2.5 Basket-Case Boundary and Initial Conditions

This subsection will focus on a discussion of the boundary conditions and the initial condition for the three-dimensional Basket-Case Model. These are just as important to the model as the PDE itself, and it is important that I discuss them and what their physical interpretation is.

My thesis focuses on numerically solving the three-dimensional version of this model, which corresponds to having three underlying assets. The choice of three underlying assets was made for two reasons. The first being that three assets is enough to explain the theory behind the Basket-Case Model. The second reason is because solving this problem with more than three assets becomes computationally infeasible with my resources.

In the three-dimensional model, there are six boundary conditions and one initial condition. The first three boundary conditions describe the case when one of the underlying assets is equal to zero for the lifetime of the option. These have already been discussed in great detail in Section 3.2.4, but I will give a short summary here for completeness. These three boundary conditions all have the same structure, so I will discuss only one of them. The others have a very similar physical meaning as well. So, the first boundary condition takes the form of

$$V(0, S_2, S_3, t) = V_1(S_2, S_3, t), \qquad \forall (S_2, S_3) \in \Omega_2 \times \Omega_3, \quad \forall t \in [0, T].$$

This boundary condition states that when the first asset is worthless for the lifetime of the option, then the value of the overall option only depends on the value of the other two underlying assets and time. This is precisely the solution to the two-dimensional model using the other two underlying assets. In other words, the value of the overall option when the first asset is worthless for the lifetime of the option is simply the value of a different Basket-Case option for the other two underlying assets.

The next three boundary conditions describe the value of the overall option when one of the underlying assets is extremely large. They too, have the same basic structure, so I will only discuss

one of them. The first of these three states

$$V(\mathbf{S}, t) \to S_1 - Ke^{-rt}, \qquad S_1 \to \infty, \quad \forall (S_2, S_3) \in \Omega_2 \times \Omega_3, \quad \forall t \in [0, T].$$

This boundary condition states that when one of the underlying assets grows very large, then the corresponding put option will almost certainly not be exercised. Therefore, by Put-Call Parity, I can derive the value of the call option to be simply the value of the asset minus the discounted strike price.

Finally, the last condition is the "initial" condition. I put this in quotes because this is really a condition at the expiration time, but with a simple change of variables (which will be done later when I begin to solve the PDE), it is transformed into an initial condition. Nonetheless, this "initial" condition takes the form of

$$V(\mathbf{S}, T) = \max\left(\|\mathbf{S}\|_1 - K, 0\right) \qquad \forall \mathbf{S} \in \Omega.$$

This condition simply gives the value of the option at expiration time, $t = T$. The reason why I am assigning a final value instead of an initial value here is because the value of the option at expiration time is much easier to calculate, as it can only be one of two things. If the sum of the values of the underlying assets is greater than the strike price, then an investor will exercise the option and its value will be the difference between these two values. Otherwise, the option will not be exercised and is worthless.

In the next section, I will state the entire Basket-Case Model for three dimensions.

3.2.6 Three-Dimensional Basket-Case Black-Scholes Model

With the PDE derived in Section 3.2.3, and the boundary conditions discussed in Section 3.2.4, I can fully state the Basket-Case Black-Scholes Model.

The Basket-Case Black-Scholes Model in three-dimensions is stated as follows:

Model 2 (Three-Dimensional Basket-Case Black-Scholes Model)**.** *Suppose that a Basket-Case call option, with strike price K and expiration time T, consists of three underlying assets whose values follows a Geometric Brownian Motion. Let their values be given by $S_1(t)$, $S_2(t)$ and $S_3(t)$. Then, let*

$$\mathbf{S} = (S_1, S_2, S_3)^T .$$

Suppose the risk-free interest rate is given by r, the volatility of the i-th asset is given by σ_i and the correlation between the i-th and j-th asset is given by ρ_{ij}. Finally, suppose that

$$\mathbf{\Omega} = \Omega_1 \times \Omega_2 \times \Omega_3,$$

where Ω_i represents the set of possible values of the i-th asset. Then the value for the call option, $V(\mathbf{S}, t)$, is modeled by

$$V_t + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) + r\nabla_S(V) \cdot \mathbf{S} - rV = 0, \tag{3.42}$$

with boundary conditions given by

$$V(0, S_2, S_3, t) = V_1(S_2, S_3, t), \qquad \forall (S_2, S_3) \in \Omega_2 \times \Omega_3, \quad \forall t \in [0, T] \tag{3.43}$$

$$V(S_1, 0, S_3, t) = V_2(S_1, S_3, t), \qquad \forall (S_1, S_3) \in \Omega_1 \times \Omega_3, \quad \forall t \in [0, T] \tag{3.44}$$

$$V(S_1, S_2, 0, t) = V_3(S_1, S_2, t), \qquad \forall (S_1, S_2) \in \Omega_1 \times \Omega_2, \quad \forall t \in [0, T] \tag{3.45}$$

$$V(\mathbf{S}, t) \to S_1 - Ke^{-rt}, \qquad S_1 \to \infty, \quad \forall (S_2, S_3) \in \Omega_2 \times \Omega_3, \quad \forall t \in [0, T] \tag{3.46}$$

$$V(\mathbf{S}, t) \to S_2 - Ke^{-rt}, \qquad S_2 \to \infty, \quad \forall (S_1, S_3) \in \Omega_1 \times \Omega_3, \quad \forall t \in [0, T] \tag{3.47}$$

$$V(\mathbf{S}, t) \to S_3 - Ke^{-rt}, \qquad S_3 \to \infty, \quad \forall (S_1, S_2) \in \Omega_1 \times \Omega_2, \quad \forall t \in [0, T] \tag{3.48}$$

$$V(\mathbf{S}, T) = \max\left(\|\mathbf{S}\|_1 - K, 0\right) \qquad \forall \mathbf{S} \in \Omega, \tag{3.49}$$

where V_i is the solution to the two-dimensional Basket-Case Model with all underlying assets except for the i-th asset.

Here I have stated the problem that I will be solving in my thesis. In my thesis, I will numerically solve this problem using the Finite Element Method (FEM).

3.3 Chapter Summary

In this chapter, I discussed the theory required to understand the Basket-Case Black-Scholes Model in three dimensions. This all lead to the main goal of this chapter, which was to state the model.

I started by introducing the one-dimensional model, sometimes called the Single Asset Model, because it helped me understand how to derive the Basket-Case Model. In this section, I talked about how delta-hedging can be thought of as a linear approximation of the value of the option. This was important, because it was a key piece in the derivation of the Single-Asset Model [8]. This derivation is sometimes called the *delta-hedging argument* because of its use of delta-hedging.

I then moved on to trying to derive the Basket-Case Model. I started by noticing that I must do some sort of modification to the models for each underlying asset in order to use Itô's Lemma. I modified these models by taking some inspiration from linear algebra and using something similar to the Gram-Schmidt Algorithm. After rewriting the models for each of the underlying assets, I used the same delta-hedging argument from the Single-Asset Model derivation to derive the Basket-Case Model.

I then took the time to discuss the interesting boundary conditions with this Model. As I discussed in Section 3.2.4, the left-side boundary conditions are actually lower-dimensional solutions to the Basket-Case Model. This was surprising, and lead to a rather interesting requirement when trying to solve the original problem: I need to solve each of the lower-dimensional problems before I can solve the original problem.

Lastly, I stated the Basket-Case Black-Scholes Model in three dimensions. The next chapter will focus on numerically solving this problem using a Finite Element Method (FEM).

Chapter 4

Numerical Solution

This chapter will focus on the numerical solving of the problem given in Section 3.2.6. For my thesis, I will be solving this problem using the Finite Element Method (FEM).

I will start by first bounding the domain. In the original model, I pose no upper limits on the domain, or the values that the underlying assets can take on. In order to solve the problem with a numerical method, I will have to bound this domain, and then solve on this bounded domain.

Next, I will quickly describe a rewriting of the model. It will be helpful for me to rewrite the model one last time in order to make the construction of the weak formulation of the PDE easier.

Then, I will talk about some of the important properties of the PDE that will justify the method I will use to solve the PDE. I will be using an Implicit-Explicit (IMEX) method to discretize the spatial domain, and thus I will need to justify the use of this method.

Finally, I will move on to the construction of the weak formulation of the PDE. This section is where I will spend the most time in this chapter because it is one of the more important aspects to the numerical solution proposed in my thesis. As mentioned above, I will be using an IMEX method to discretize the spatial domain, and here I will explain in detail the particular discretization I will be using.

At the end of this chapter, I will have the weak formulation for my PDE. This weak formulation will then be used in a `C++` program that I have written, using a Finite Element Library known as `deal.II`, to solve the PDE [5].

4.1 Truncation of the Spatial Domain

Before solving the Basket-Case Model, I need to truncate the spatial domain. The original model proposed in Section 3.2.6 did not impose any upper limits on the spatial domain, so in this section I will discuss what restrictions I will use to truncate the domain.

Recall that my spatial domain is simply the tensor product of the domains for each underlying asset. Therefore, it will be sufficient to bound the domain for each underlying asset.

There are two values then, that I need to prescribe in order to bound an asset's domain. These values are the minimum and maximum values that each asset will obtain. Because the minimum value is easier to understand, I will prescribe this value first.

Realistically, the asset's lowest value is achieved when it is worthless. This means that the lowest value that I will let each asset achieve in my model is a value of zero. So here, I have

$$0 \leq S_i(t), \quad \forall i \in \{1, 2, 3\}. \tag{4.1}$$

The upper bound for each asset's value is a bit trickier to prescribe. In reality, there is no limit to how valuable an asset can be. I therefore need to choose a value that is "big enough" in order for me to still be able to understand the solution to the PDE and how it behaves. This "big enough" value is for me to decide, but it should be something that will allow me to see *mostly* how the solution to the PDE behaves. I will, therefore, look at the boundary conditions for my model to see how the solution to the PDE should behave as an asset's value gets very large. Recall, that this condition takes the form of

$$V(\mathbf{S}, t) \to \infty, \quad S_i \to \infty.$$

This equation holds for all assets, so I only need to consider one of them to understand the rest. Now, with this condition, I see that the value of the option simply grows very large as the value of one of the underlying assets grows very large. Notice that this tells me that the value of the option doesn't really depend on how large the strike price, K, is when the value of the underlying asset gets "big enough". This is intuitively because the cost of the strike price becomes insignificant after a certain point. This means that any increases in the value of the asset after this "big enough" value, won't make a large difference in how the solution to the PDE behaves. So, to prescribe an upper bound for the value of the underlying assets, I will choose an asset value that is on the order

of 10 times the strike price. This value was chosen empirically by increasing the value of the upper bound of the asset value until no differences in the solution were found. This condition gives me the following

$$S_i(t) \leq 10K, \quad \forall i \in \{1, 2, 3\}. \tag{4.2}$$

Putting Equations (4.1) and (4.2) together, I arrive at the following bound for the values of each underlying asset

$$0 \leq S_i(t) \leq 10K, \quad \forall i \in \{1, 2, 3\}. \tag{4.3}$$

To construct the entire bounded domain, I will simply have the tensor product of all three of these bounded domains.

These bounded domains are created to allow me to get an *understanding* of the solution to the PDE. In my thesis, I am not actually concerned with the value of the solution at any time and asset price, but I am more concerned with building an understanding of how the solution behaves. If I wanted to know the value of the option when one of the assets had a value outside of this bounded domain that I have constructed, then I could simply increase the size of the domain to include this value. Since I am concerned with the behavior of the solution, then this domain will be sufficient for my thesis.

In the next section, I will describe one final reformulation of my model so that the weak formulation is easier to construct.

4.2 Conservative Formulation

This section will focus on the reformulation of my model into what is known as the *Conservative Formulation*. Recall that the Basket-Case PDE that I stated in Section 3.2.6 was

$$V_t + \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) + r\nabla_S(V) \cdot \mathbf{S} - rV = 0. \tag{4.4}$$

This form of the PDE is known as the *non-conservative* form. Unfortunately, in this form, the weak formulation is not very easy to construct. If I reformulate the PDE into it's conservative form, then the weak formulation construction will be much easier.

The biggest reason for my desire to change to the conservative formulation is because of the

$$\frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right) : \nabla_S^2(V) \tag{4.5}$$

term, which will be difficult to integrate by parts when I am constructing the weak formulation.

To start, I will give a lemma that will help me rewrite the term in Equation (4.5). This lemma is a special case of the product rule for vector valued functions.

Lemma 11. *Suppose that* $\mathbf{M}(\mathbf{S})$ *is a matrix that is a function of the vector* $\mathbf{S}$*. Suppose also that* $V(\mathbf{S})$ *is a function of the vector* $\mathbf{S}$*. Then,*

$$\nabla_S \cdot (\mathbf{M}[\nabla_S V]) = (\nabla_S \cdot \mathbf{M}) \cdot \nabla_S V + \mathbf{M} : \nabla_S^2 V. \tag{4.6}$$

Here,

$\nabla_S \cdot (\mathbf{M}[\nabla_S V])$ *represents the divergence of the vector* $\mathbf{M}[\nabla_S V]$,

$\nabla_S \cdot \mathbf{M}$ *represents the divergence of the matrix* $\mathbf{M}$,

and $\nabla_S^2 V$ *represents the Hessian of the function V.*

The proof of this lemma took inspiration from the proof of the multi-dimensional stochastic product rule. This is seen through the use of Einstein notation in this proof to rewrite the left-hand side of the equation. I will simply expand the term on the right using Einstein Notation to show that it equals the right-hand side.

Proof. Using Einstein Notation, I can rewrite $\nabla_S \cdot (\mathbf{M}[\nabla_S V])$ as

$$\begin{aligned}\nabla_S \cdot (\mathbf{M}[\nabla_S V]) &= \frac{\partial}{\partial S_i}\left(M_{ij}\left(\nabla_S V\right)_j\right) \\ &= \frac{\partial}{\partial S_i}\left(M_{ij}\right)\frac{\partial V}{\partial S_j} + M_{ij}\frac{\partial^2 V}{\partial S_i \partial S_j}.\end{aligned}$$

At this point, I have simply expanded the left-hand side into its sum representation, and then applied the "normal" product rule. I am allowed to do this, because each term is simply a partial derivative of a *real-valued* function, instead of the original vector-valued function I started trying to differentiate. Next, I notice that the term

$$\frac{\partial}{\partial S_i}\left(M_{ij}\right)$$

is something very special. This term, by definition, is the j-th term of the *divergence* of a matrix. So, I see that I can rewrite the equation as follows

$$\nabla_S \cdot (\mathbf{M}[\nabla_S V]) = \left(\nabla_S \cdot \mathbf{M}\right)_j \left(\nabla_S V\right)_j + M_{ij}\left(\nabla_S^2 V\right)_{ij}.$$

Here, I almost have the right-hand side that I am looking for. The only thing that I need to do is deal with the last remaining terms. This is done by interpreting the sums as their vector operations to leave me with

$$\begin{aligned}\nabla_S \cdot (\mathbf{M}[\nabla_S V]) &= \left(\nabla_S \cdot \mathbf{M}\right)_j \left(\nabla_S V\right)_j + M_{ij}\left(\nabla_S^2 V\right)_{ij} \\ &= \left(\nabla_S \cdot \mathbf{M}\right) \cdot \nabla_S V + \mathbf{M} : \nabla_S^2 V.\end{aligned}$$

This then, completes the proof. □

I am now ready to rewrite the Basket-Case PDE into its conservative form. I will do this by using Lemma 11 to rewrite the "difficult" term in Equation (4.5). First, however, I will perform the

following minor substitution to make the notation easier

$$\mathbf{A} \coloneqq \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right). \tag{4.7}$$

With this substitution, the Basket-Case PDE becomes

$$V_t + \mathbf{A} : \nabla_S^2(V) + r\nabla_S(V) \cdot \mathbf{S} - rV = 0. \tag{4.8}$$

I see here, that the second term in Equation (4.8) resembles one of the terms in Lemma 11. More specifically, by applying Lemma 11 and rearranging some terms, I see that

$$\mathbf{A} : \nabla_S^2 V = \nabla_S \cdot (\mathbf{A}\left[\nabla_S V\right]) - (\nabla_S \cdot \mathbf{A}) \cdot \nabla_S V. \tag{4.9}$$

So, substituting this into Equation (4.8) and combining like terms, I finally arrive at the Conservative Formulation for the Basket-Case PDE:

$$V_t + \nabla_S \cdot (\mathbf{A}\left[\nabla_S V\right]) + \left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A})\right] \cdot \nabla_S V - rV = 0. \tag{4.10}$$

In this section, I have rewritten the Basket-Case PDE from its Non-Conservative Formulation into its Conservative-Formulation. This new formulation, seen in Equation (4.10), may appear more complex, but will be much easier to deal with in the next section when I am constructing the weak formulation. The reason is because the conservative form will be much easier to integrate by parts.

I will use the conservative form in the following section to construct the weak formualtion for the Basket-Case Black-Scholes PDE.

4.3 Weak Formulation

This section will focus on constructing the weak formulation of the PDE stated in Equation (4.10). Constructing the weak formulation is a crucial step in generating a numerical solution using the Finite Element Method, so I will go into great detail explaining this construction.

I will first explain the need for a slight change of variables in order to deal with the "final" condition that I have for the Basket-Case Model. Then, I will describe how to use the weak formulation to generate a system of linear equations that I will use to generate a solution to the PDE. This part will be done in several subsections because I want to take the time to describe all of the parts. After this, I will state the system of linear equations that are going to be solved. Finally, I will give a bit of theory justifying a few of my decisions when constructing the weak formulation. These will have to do with the convergence of the numerical solution.

4.3.1 Time Reversal

The Basket-Case Model stated in Section 3.2.6 has a somewhat peculiar condition. Instead of an initial condition, I have a *final* condition. This means that I have a condition on the value of the option at expiration time, instead of at the initial time. In order to correct for this, I will make the following change of variables

$$\tau = T - t. \tag{4.11}$$

This change of variables simply *reverses* the time, which will let me solve the PDE in reverse.

So, with this change of variables, the Basket-Case PDE becomes

$$V_\tau - \nabla_S \cdot (\mathbf{A}\left[\nabla_S V\right]) - \left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A})\right] \cdot \nabla_S V + rV = 0. \tag{4.12}$$

This PDE will now be the one that I will actually generate a weak formulation for. With the time reversal that I have done, the final condition now becomes a true initial condition. In the next section, I will begin the weak formulation of this PDE. This process begins with the discretization of the time domain.

4.3.2 Time Domain Discretization

In order to construct a weak formulation to the Basket-Case PDE, I will first discretize my time domain. This process is important because it allows me to construct solutions on small pieces of the time domain, and then "stitch" these solutions back together at the end to create the complete solution. For my numerical solution, I will do this by partitioning up the time domain into N_t evenly sized intervals. This will result in the following partition:

$$[0, T] = [\tau_0 = 0, \tau_1) \cup [\tau_1, \tau_2) \cup \cdots \cup [\tau_{N_t-2}, \tau_{N_t-1} = T). \tag{4.13}$$

It is also worth noting that in this partition, each interval has a width of

$$k = \frac{T}{N_t}. \tag{4.14}$$

I will call this value, k, the *step-size*. In my thesis, this value will remain constant. It should be noted, however, that this value does not have to be constant. I am choosing to use a constant value here because my thesis is focused on generating a numerical solution to the Basket-Case Model, not optimizing the calculation of this numerical solution.

Since the step-size is constant, I can express the left-hand side of each interval as

$$\tau_n = nk, \qquad n \in \{0, \ldots, N_t - 1\}. \tag{4.15}$$

This is helpful because I can then discuss the value of the option at each of these points. Instead of trying to get the value at each point in time, I will simply calculate the value at each of the points in Equation (4.15). This is the first step in creating the system of linear equations that I will ultimately solve. So, I will let

$$V^n(\mathbf{S}) = V(\mathbf{S}, \tau_n). \tag{4.16}$$

Next, I need to approximate the time derivative using this discretization. I will do this using a first order explicit approximation of the form

$$V_t \approx \frac{V^n(\mathbf{S}) - V^{n-1}(\mathbf{S})}{k} \tag{4.17}$$

Not only do I need to rewrite the time derivative using this time discretization, but I also need to rewrite all of the other terms in the PDE. In order to accomplish this, I will use a *theta-scheme*. This method has been used to solve diffusion-reaction equations because of its usefulness in splitting the diffusion and non-diffusion terms [3] [18]. This method allows for me to solve the diffusive terms in a separate way than the non-diffusive terms. The PDE I am solving in my thesis is a diffusive-advective-reactive PDE, but this method is still useful because of the ability to split the diffusive terms from the non-diffusive terms.

So, using Equation (4.17) to rewrite the time derivative, and the theta-method to rewrite the other terms, I have

$$\begin{aligned} 0 = & \frac{V^n - V^{n-1}}{k} \\ & - (1-\theta_B)\,\nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^{n-1}\right]\right) - \theta_B \nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) \\ & - (1-\theta_C)\left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A})\right] \cdot \nabla_S V^{n-1} - \theta_C \left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A})\right] \cdot \nabla_S V^n \\ & + (1-\theta_M)\, r V^{n-1} + \theta_C r V^n. \end{aligned} \tag{4.18}$$

Here, I have suppressed the dependence on $\mathbf{S}$, and I have also moved the non-zero terms to the right side of the equation for clarity. Notice also, that I have allowed myself the liberty of assigning a different theta value to each term. I will discuss in the following sections what exactly these thetas should be, but I want to emphasize that I will be using different values depending on the properties of each term.

With Equation (4.18) constructed, I have finished the time-discretization of the Basket-Case PDE. In order to arrive at this equation I had to do several things. I first had to partition my time

domain into evenly spaced intervals. Then, I approximated the time-derivative using a first order approximation. Finally, I had to use a theta-method to rewrite the other terms in the PDE.

Next, I will perform the spatial discretization. A few more interesting things will happen there, because I will have to deal with the diffusive, advective and reactive terms all separately since they all have slightly different properties.

4.3.3 Spatial Discretization

The time discretization is only half of the creation of the weak formulation. The other half is the spatial discretization.

For this, I will consider partitioning each of the asset price domains into N_s equally sized intervals. Because my spatial domain is the tensor product of three of these asset price domains, then this process will result in $(N_s)^3$ cubes.

At this point, I will define a test function that is piece-wise tri-linear on each of these cubes. This will result in the the set

$$\left\{\phi_i(\mathbf{S}) \;\middle|\; 1 \leq i \leq (N_s)^3\right\} \tag{4.19}$$

of test functions. These will be considered a set of *basis functions* for my space of solutions. What this means, is that I will approximate my actual solution at each time-step, V^n, by another function, V_h^n, that can be expressed as

$$V_h^n(\mathbf{S}) = \sum_{j=1}^{(N_s)^3} V_j^n \phi_j(\mathbf{S}). \tag{4.20}$$

Expressing the numerical solution like this means that I really only need to solve for the coefficients V_j^n. How exactly I will solve for these coefficients is the major topic of this section. This idea, however, is the main idea of the Finite Element Method, and it is how I am ultimately going to construct a numerical solution to the Basket-Case Model.

So, in order to solve for these coefficients, I am going to construct the weak formulation of the PDE. To do this, I will simply multiply Equation (4.18) by each test function, ϕ_j, and then integrate over the domain, Ω. For convenience, I will also multiply this equation by k in order to get rid of

fractions. Doing this, I arrive at the following equation:

$$0 = \int_\Omega \phi_i \left[V^n - V^{n-1}\right] d\mathbf{S} \tag{4.21}$$

$$- k \int_\Omega \phi_i \left[(1-\theta_B)\nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^{n-1}\right]\right) - \theta_B \nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right)\right] d\mathbf{S} \tag{4.22}$$

$$- k \int_\Omega \phi_i \left[(1-\theta_C)\left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A})\right] \cdot \nabla_S V^{n-1} - \theta_C \left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A})\right] \cdot \nabla_S V^n\right] d\mathbf{S} \tag{4.23}$$

$$+ k \int_\Omega \phi_i \left[(1-\theta_M) r V^{n-1} + \theta_C r V^n\right] d\mathbf{S}. \tag{4.24}$$

I have suppressed the dependencies on $\mathbf{S}$ in this equation for clarity. Also, note that this equation must hold for *all* test functions in the set given in Equation (4.19).

This equation is quite complex and has several parts to it, so I will evaluate each part separately in the following subsections. What is important to see here is that, when I eventually substitute Equation (4.20) for the solution at each time-step, I will have a system of equations with $(N_s)^3$ unknowns and $(N_s)^3$ equations. This will be the system of linear equations that I will ultimately solve in order to get the numerical solution.

First Integral

The first integral I am going to examine is given in Equation (4.21). This integral has two parts that have the same general form

$$\int_\Omega \phi_i V^n d\mathbf{S}. \tag{4.25}$$

There are some minor differences between the two parts, but I will account for these at the end of this section.

First, I will approximate V^n by Equation (4.20), to rewrite the integral as

$$\int_\Omega \phi_i V^n d\mathbf{S} \approx \int_\Omega \phi_i \sum_{j=1}^{(N_s)^3} V_j^n \phi_j d\mathbf{S}.$$

Recall that each V_j^n is simply a scalar, and that I am trying to solve for these scalars.

Now, I can pull out the sum, since it is a finite sum, and rearrange some terms to get

$$\int_\Omega \phi_i \sum_{j=1} V_j^n \phi_j d\mathbf{S} = \sum_j V_j^n \int_\Omega \phi_i \phi_j d\mathbf{S}.$$

Here, I have taken the liberty of suppressing the limits of summation for clarity.

Next, I will use the following definition to further simplify the integral.

Definition 9. Suppose that f and g are integrable functions over domain Ω. Then, define

$$(f, g) \coloneqq \int_\Omega f g dx.$$

Using this definition, I can rewrite the integral term and get

$$\begin{aligned}\int_\Omega \phi_i \sum_{j=1} V_j^n \phi_j d\mathbf{S} &= \sum_j V_j^n \left(\phi_i, \phi_j\right) \\ &= \sum_j V_j^n M_{ij},\end{aligned}$$

where $M_{ij} \coloneqq (\phi_i, \phi_j)$. Recall that this equation must hold for *all* test functions, ϕ_i. Therefore, this integral leads to the following set of linear equations:

$$\mathbf{M}\mathbf{V}^n. \tag{4.26}$$

Finally, going back to the original integral in Equation (4.21), this integral leads to the following set of linear equations:

$$\mathbf{M}\mathbf{V}^n - \mathbf{M}\mathbf{V}^{n-1}. \tag{4.27}$$

It is worth noting that in the literature, this matrix $\mathbf{M}$, is often referred to as the *mass matrix* [21].

In this section, I started with a system of $(N_s)^3$ equations in the form of integrals and rewrote them in terms of a matrix multiplied by a vector. This was accomplished by using Equation (4.20)

to approximate the solution to the PDE, and then grouping terms together to arrive at the matrix-vector product. With this representation, I will be solving for the coefficients that correspond to each basis function. With these coefficients then, I will be able to construct the numerical solution to the PDE.

I only need to construct these matrix-vector products for the other integrals.

Second Integral

The next integral I am going to examine is given in Equation (4.22). This integral also has two parts that have the same general form

$$-\int_{\Omega} \phi_i \nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) d\mathbf{S}. \tag{4.28}$$

As with the previous section, I will examine this integral and account for the differences at the end.

One difference between this integral and the previous one is that here, I will have second-derivative terms for the solution in this integral. For this reason, I cannot simply replace V^n with its approximation given in Equation (4.20). Recall, that this approximation has "kinks", and therefore I cannot have second-derivative terms. Therefore, I will use integration-by-parts to get rid of this second-derivative term. Doing this, I will rewrite the integral as

$$-\int_{\Omega} \phi_i \nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) d\mathbf{S} = -\int_{\partial\Omega} \phi_i \left(\mathbf{A}\left[\nabla_S V^n\right]\right) \cdot d\mathbf{S} + \int_{\Omega} \nabla_S \phi_i \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) d\mathbf{S}. \tag{4.29}$$

At this point, I will use one last property of these test functions: they are all zero at the boundary of the domain. This means then, that Equation (4.29) will simply become

$$-\int_{\Omega} \phi_i \nabla_S \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) d\mathbf{S} = \int_{\Omega} \nabla_S \phi_i \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) d\mathbf{S}. \tag{4.30}$$

In this equation, there are no more second-derivative terms for the solution. I am allowed, therefore, to replace V^n with its approximation. So, I will follow the same pattern I followed in

Section 4.3.3 to derive the matrix-vector product version of these system of equations. I will go ahead and take the liberty of suppressing the limits of summation when using the approximation for V^n for clarity. So,

$$\begin{aligned}\int_\Omega \nabla_S \phi_i \cdot \left(\mathbf{A}\left[\nabla_S V^n\right]\right) d\mathbf{S} &= \int_\Omega \nabla_S \phi_i \cdot \left(\mathbf{A}\left[\nabla_S \left(\sum_j V_j^n \phi_j\right)\right]\right) d\mathbf{S} \\ &= \int_\Omega \nabla_S \phi_i \cdot \left(\sum_j V_j^n \mathbf{A}\left[\nabla_S \phi_j\right]\right) d\mathbf{S} \\ &= \sum_j V_j^n \int_\Omega \nabla_S \phi_i \cdot \left(\mathbf{A}\left[\nabla_S \phi_j\right]\right) d\mathbf{S} \\ &= \sum_j V_j^n \left(\nabla_S \phi_i, \mathbf{A}\left[\nabla_S \phi_j\right]\right).\end{aligned}$$

Here again, just as in Section 4.3.3, I will create a matrix from the inner product term. Therefore, I will let

$$B_{ij} = \left(\nabla_S \phi_i, \mathbf{A}\left[\nabla_S \phi_j\right]\right), \tag{4.31}$$

which will make the system of linear equations become

$$\mathbf{B}\mathbf{V}^n. \tag{4.32}$$

Accounting for the differences between the general integral and the real integral in Equation (4.22), I arrive at the following system of linear equations:

$$k\left(1-\theta_B\right)\mathbf{B}\mathbf{V}^{n-1} + k\theta_B \mathbf{B}\mathbf{V}^n. \tag{4.33}$$

In this section, I derived the matrix-vector form of the weak formulation. The main point that I want to highlight in this section is the use of integration-by-parts. If I had not written the Basket-Case PDE into its conservative form, then this step would have been very difficult. The integration-by-parts step allowed me to substitute for the approximation of the solution, which allowed me to derive the matrix-vector form of the weak formulation.

Third Integral

The third integral I am going to examine is given in Equation (4.23). As before, this integral has 2 parts and I will consider a general form of the integral. This general form looks like

$$-\int_{\Omega} \phi_i \left[r\mathbf{S} - (\nabla_S \cdot \mathbf{A}) \right] \cdot \nabla_S V^n d\mathbf{S}. \tag{4.34}$$

For simplicity, I will go ahead and distribute the negative sign in front of the integral to the terms inside. This results in

$$\int_{\Omega} \phi_i \left[(\nabla_S \cdot \mathbf{A}) - r\mathbf{S} \right] \cdot \nabla_S V^n d\mathbf{S}. \tag{4.35}$$

In the above integral, I have no second-derivative terms for V^n, so I am allowed to substitute V^n for its approximation. Doing this is very similar to the previous sections, so I will be a bit more brief. Doing the substitution and grouping terms together, I have

$$\begin{aligned}
\int_{\Omega} \phi_i \left[(\nabla_S \cdot \mathbf{A}) - r\mathbf{S} \right] \cdot \nabla_S V^n d\mathbf{S} &= \int_{\Omega} \phi_i \left[(\nabla_S \cdot \mathbf{A}) - r\mathbf{S} \right] \cdot \nabla_S \left(\sum_j V_j^n \phi_j \right) d\mathbf{S} \\
&= \sum_j V_j^n \int_{\Omega} \left(\phi_i \left[(\nabla_S \cdot \mathbf{A}) - r\mathbf{S} \right], \nabla_S \phi_j \right) \\
&= \sum_j V_j^n C_{ij} \\
&= \mathbf{C}\mathbf{V}^n,
\end{aligned}$$

where I defined the matrix, $\mathbf{C}$, by

$$C_{ij} \coloneqq \left(\phi_i \left[(\nabla_S \cdot \mathbf{A}) - r\mathbf{S} \right], \nabla_S \phi_j \right).$$

Finally, after accounting for the differences between the general integral and the integral in Equation (4.23), I have the following system of equations represented with matrix-vector multipli-

cation:

$$k\left(1-\theta_C\right)\mathbf{CV}^{n-1}+k\theta_C\mathbf{CV}^{n}. \tag{4.36}$$

This system of linear equations was not much different than the previous two, so I took the liberty of only highlighting the important steps. The key here again, was to substitute for the approximation and group terms.

Fourth Integral

The final integral I need to examine is given in Equation (4.24). As I have done several times already, I will first consider the general form of this integral:

$$\int_{\Omega}\phi_i V^n d\mathbf{S}. \tag{4.37}$$

This integral is exactly the general integral considered in Section 4.3.3, so I will directly substitute the results from that section to obtain the following system of equations expressed as a matrix-vector product:

$$kr\left(1-\theta_M\right)\mathbf{MV}^{n-1}+kr\theta_M\mathbf{MV}^{n}. \tag{4.38}$$

With this system of equations represented as a matrix-vector product, I have completed the spatial discretization of the Basket-Case PDE. In the past four sections, I have taken each set of integral equations and expressed them as matrix-vector products. The importance of doing this, is that these equations become much simpler to solve for each V_j^n. These coefficients are what I am trying to solve for as well, because with them, I can reconstruct the numerical solution.

In the next section, I will put the past four sections together to build the full weak formulation model.

4.4 Complete Model

In the previous section, I outlined the spatial discretization. This section will focus on taking the results from Section 4.3.3 and compiling them into one complete weak formulation. Once I

have the complete system of equations that need to be solved, I can use a computer program to calculate the individual coefficients, V_j^n.

Putting together the results from the previous section, I have the following weak formulation:

$$\begin{aligned} 0 = &\mathbf{M}\mathbf{V}^n - \mathbf{M}\mathbf{V}^{n-1} \\ &+ k\left(1-\theta_B\right)\mathbf{B}\mathbf{V}^{n-1} + k\theta_B\mathbf{B}\mathbf{V}^n \\ &+ k\left(1-\theta_C\right)\mathbf{C}\mathbf{V}^{n-1} + k\theta_C\mathbf{C}\mathbf{V}^n \\ &+ kr\left(1-\theta_M\right)\mathbf{M}\mathbf{V}^{n-1} + kr\theta_M\mathbf{M}\mathbf{V}^n. \end{aligned} \tag{4.39}$$

This equation is the result of the time and spatial discretizations that I have done up to this point. I am almost ready to solve this equation for the coefficients, but I have one more problem I need to work out. I need to determine what each θ should be equal to. I stated at the beginning of this chapter that I would use a theta-method in the time-discretization, but I would decide exactly what values of θ to use later based on the properties of the problem.

I am going to ultimately want a system of equations that I can use to solve for $\mathbf{V}^n$, given that I already know $\mathbf{V}^{n-1}$. The method for solving the linear system of equations I am going to use is the Conjugate Gradient Method. For this method to work, however, I need to make sure that the matrix that is being multiplied by $\mathbf{V}^n$ is Symmetric Positive Definite (SPD). So, I will choose each θ to give me stability along with convergence.

First, I will look at the matrix $\mathbf{C}$. Recall that this matrix is given by

$$C_{ij} \coloneqq \left(\phi_i\left[\left(\nabla_S \cdot \mathbf{A}\right) - r\mathbf{S}\right], \nabla_S\phi_j\right).$$

Because this matrix is not symmetric, it will not be allowed to be multiplied by $\mathbf{V}^n$. In other words, I will have to treat these terms in an *explicit* manner. So, I will choose

$$\theta_C = 0. \tag{4.40}$$

Next, I will look at the matrix $\mathbf{B}$. This matrix, as it turns out, is Symmetric Positive Semi-Definite (SPSD). To show this, I will first take a look at the matrix $\mathbf{A}$.

Lemma 12. *The matrix* $\mathbf{A}$ *as given in Equation* (4.7) *is SPSD.*

Proof. Recall that $\mathbf{A}$ is given by

$$\mathbf{A} = \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right).$$

To show symmetry, I need to show that $\mathbf{A}^T = \mathbf{A}$. So,

$$\begin{aligned}
\mathbf{A}^T &= \left(\frac{1}{2}\mathbf{G}\mathbf{G}^T\right)^T \\
&= \frac{1}{2}\left(\mathbf{G}\mathbf{G}^T\right)^T \\
&= \frac{1}{2}\mathbf{G}\mathbf{G}^T \\
&= \mathbf{A}
\end{aligned}$$

Therefore, the matrix $\mathbf{A}$ is symmetric.

To show positive semi-definiteness, I need to show

$$\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0, \qquad \forall \mathbf{x} \neq 0.$$

So, let $\mathbf{x}$ be a non-zero vector in $\mathbb{R}^3$. Then,

$$\begin{aligned}
\mathbf{x}^T\mathbf{A}\mathbf{x} &= \mathbf{x}^T\left(\frac{1}{2}\mathbf{G}\mathbf{G}^T\right)\mathbf{x} \\
&= \frac{1}{2}\left(\mathbf{G}^T\mathbf{x}\right)^T\left(\mathbf{G}^T\mathbf{x}\right) \\
&= \frac{1}{2}\|\mathbf{G}^T\mathbf{x}\|_2 \\
&\geq 0.
\end{aligned}$$

Thus, the matrix $\mathbf{A}$ is SPSD. □

I needed to show that the matrix $\mathbf{A}$ was SPSD because this matrix is an important component in the matrix $\mathbf{B}$. My real goal is to show that the matrix $\mathbf{B}$ is SPSD, but I first needed to show this for the matrix $\mathbf{A}$. I want this SPSD property for the matrix $\mathbf{B}$ because then I will be able to use the Conjugate Gradient Method to solve for terms that include $\mathbf{B}$. So, I will now show that the matrix $\mathbf{B}$ is SPSD.

Lemma 13. *Suppose the matrix* $\mathbf{B}$ *is given by*

$$B_{ij} = \left(\nabla_S \phi_i, \mathbf{A}\left[\nabla_S \phi_j\right]\right).$$

Then this matrix is SPSD.

Proof. I want to show that the matrix $\mathbf{B}$ is both symmetric and positive semi-definite. So, I will first show the symmetry of this matrix. To do this, I will show that $B_{ij} = B_{ji}$. So,

$$\begin{aligned} B_{ij} &= \left(\nabla_S \phi_i, \mathbf{A}\left[\nabla_S \phi_j\right]\right) \\ &= \int_\Omega \nabla_S \phi_i \cdot \mathbf{A} \nabla_S \phi_j d\mathbf{S} \\ &= \int_\Omega \frac{\partial \phi_i}{\partial S_k} \left(\mathbf{A} \nabla_S \phi_j\right)_k d\mathbf{S}, \end{aligned}$$

where I expanded the dot product using Einstein Notation in the last line. Continuing the expansion, I have

$$\begin{aligned} B_{ij} &= \int_\Omega \frac{\partial \phi_i}{\partial S_k} \left(\mathbf{A} \nabla_S \phi_j\right)_k d\mathbf{S} \\ &= \int_\Omega \frac{\partial \phi_i}{\partial S_k} A_{kr} \frac{\partial \phi_j}{\partial S_r} d\mathbf{S}. \end{aligned}$$

Now, at this point, I will use the result from Lemma 12. More specifically, I will use the fact that the matrix $\mathbf{A}$ is symmetric. Thus,

$$\begin{aligned} B_{ij} &= \int_{\Omega} \frac{\partial \phi_i}{\partial S_k} A_{kr} \frac{\partial \phi_j}{\partial S_r} d\mathbf{S} \\ &= \int_{\Omega} \frac{\partial \phi_i}{\partial S_k} A_{rk} \frac{\partial \phi_j}{\partial S_r} d\mathbf{S} \\ &= \int_{\Omega} \frac{\partial \phi_j}{\partial S_r} \left(\mathbf{A} \nabla_S \phi_j\right)_r d\mathbf{S} \\ &= \int_{\Omega} \nabla_S \phi_j \cdot \mathbf{A} \nabla_S \phi_i d\mathbf{S} \\ &= B_{ji}. \end{aligned}$$

Thus, the matrix $\mathbf{B}$ is symmetric because $B_{ij} = B_{ji}$. The next step, then, is to show the positive semi-definiteness property. For this, I will consider a non-zero vector in $\mathbb{R}^n$, say $\mathbf{x}$. Then,

$$\mathbf{x}^T \mathbf{B} \mathbf{x} = x_i \left(\mathbf{B}\mathbf{x}\right)_i,$$

where I have used Einstein Notation again to expand things. Continuing, I have

$$\begin{aligned} \mathbf{x}^T \mathbf{B} \mathbf{x} &= x_i B_{ij} x_j \\ &= x_i \left(\int_{\Omega} \nabla_S \phi_i \cdot \mathbf{A} \nabla_S \phi_j d\mathbf{S} \right) x_j \\ &= \int_{\Omega} x_i \frac{\partial \phi_i}{\partial S_k} A_{kr} x_j \frac{\partial \phi_j}{\partial S_r} d\mathbf{S}. \end{aligned}$$

At this point, I will make the following substitution

$$u_k \coloneqq x_i \frac{\partial \phi_i}{\partial S_k}. \tag{4.41}$$

The reason is because, with this substitution, some of the terms collect nicely:

$$\begin{aligned}\mathbf{x}^T\mathbf{B}\mathbf{x} &= \int_\Omega x_i \frac{\partial \phi_i}{\partial S_k} A_{kr} x_j \frac{\partial \phi_j}{\partial S_r} d\mathbf{S} \\ &= \int_\Omega u_k A_{kr} u_r d\mathbf{S} \\ &= \int_\Omega \mathbf{u}^T \mathbf{A}\mathbf{u} d\mathbf{S}.\end{aligned}$$

At this point, I will use the other half of the result from Lemma 12. Because the matrix $\mathbf{A}$ is positive semi-definite, I know that the integrand is never negative. Therefore, the overall integral will never be negative either. So, I must have

$$\begin{aligned}\mathbf{x}^T\mathbf{B}\mathbf{x} &= \int_\Omega \mathbf{u}^T \mathbf{A}\mathbf{u} d\mathbf{S} \\ &\geq 0.\end{aligned}$$

With this, I have proved that the matrix $\mathbf{B}$ is SPSD. □

In fact, it comes as no surprise that the matrix $\mathbf{B}$ is SPSD. The reason is because this term corresponds to the *diffusive* term in the Basket-Case PDE. Therefore, it should have this property. Furthermore, because this term corresponds to the diffusive part of the PDE, I will want to solve it with some sort of implicit scheme. Therefore, because of its good convergence rate, I will use the Crank-Nicolson Scheme for the terms corresponding to the matrix $\mathbf{B}$ [22]. This corresponds to

$$\theta_B = \frac{1}{2}. \tag{4.42}$$

Finally, I simply need to choose a value for θ_M. For this, I will choose

$$\theta_M = \frac{1}{2} \tag{4.43}$$

as well in order to keep the $\mathcal{O}(k^2)$ convergence that I get from the Crank-Nicolson Method. So, substituting these values for my various thetas, Equation (4.39) I get:

$$\begin{aligned} 0 &= \mathbf{M}\mathbf{V}^n - \mathbf{M}\mathbf{V}^{n-1} \\ &+ k\frac{1}{2}\mathbf{B}\mathbf{V}^{n-1} + k\frac{1}{2}\mathbf{B}\mathbf{V}^n \\ &+ k\mathbf{C}\mathbf{V}^{n-1} \\ &+ kr\frac{1}{2}\mathbf{M}\mathbf{V}^{n-1} + kr\frac{1}{2}\mathbf{M}\mathbf{V}^n. \end{aligned} \tag{4.44}$$

Finally, grouping terms together and moving all of the $\mathbf{V}^n$ terms to the left side of the equation, I have

$$\left(\left[1 + \frac{1}{2}kr\right]\mathbf{M} + \frac{1}{2}k\mathbf{B}\right)\mathbf{V}^n = \left(\left[1 - \frac{1}{2}kr\right]\mathbf{M} - \frac{1}{2}k\mathbf{B} - k\mathbf{C}\right)\mathbf{V}^{n-1}. \tag{4.45}$$

This, finally, is the fully-discrete formulation for the Basket-Case Black-Scholes Model that I will be using to solve the PDE.

Before I close this chapter, I should note one more thing about how this will actually be solved. In order to solve for the coefficients at the n-th time-step, I will have to invert the matrix

$$\left[1 + \frac{1}{2}kr\right]\mathbf{M} + \frac{1}{2}k\mathbf{B}. \tag{4.46}$$

The need to invert this matrix is most of the reason that I went through the effort to prove Lemma 13. Because the mass matrix, $\mathbf{M}$, is SPD, then the matrix in Equation (4.46) *must* be SPD as well. This is because, along with k and r being positive values, the matrix $\mathbf{B}$ is SPSD. Thus, since this matrix is SPD, I can use the Conjugate Gradient Method to invert this matrix and solve for the coefficients.

Recall that in the construction of the weak formulation in Equation (4.45), I decided to handle the diffusive terms differently than the other terms. In particular, I handled them differently than the advective term, which corresponds to the term with the matrix $\mathbf{C}$. In some cases, this may lead to instability if the advective terms are too "important" relative to the diffusive terms [16]. In the

next section, I will discuss this topic slightly in order to give an understanding of why I did not include any extra terms to handle instability.

4.5 Peclet Number

This section will focus on a value called the *Peclet Number*. Because the Basket-Case PDE has an advection and a diffusive term, there may be instability in solving for the numerical solution using the weak formulation I proposed in Equation (4.45). In some cases, some extra terms are needed to have stability, but as I will give evidence for in this section, this is not the case in the numerical scheme I have presented in this thesis.

The Peclet Number, as it concerns to my thesis, is a number that relates the advection of the value of the option to the diffusion of the value of the option. It is a dimensionless value, and is defined on each cell of the mesh. Given a uniform mesh with each cell having a diameter of h, the Peclet number is given by:

$$P_{\text{cell}}\left(\mathbf{S}, h\right) = \frac{\beta(\mathbf{S})h}{\kappa(\mathbf{S})}. \tag{4.47}$$

Here,

$$\beta(\mathbf{S}) = \|r\mathbf{S} - \nabla \cdot \mathbf{A}(\mathbf{S})\|_2 \qquad \text{and,} \tag{4.48}$$

$$\kappa(\mathbf{S}) = \|\mathbf{A}(\mathbf{S})\|_2. \tag{4.49}$$

Physically, this value is trying to find the ratio of the "advective-ness" of the PDE to the "diffusive-ness" of it. The β term corresponds to the amount of this "advective-ness", while the κ term corresponds to the amount of "diffusive-ness".

If $P_{\text{cell}}\left(\mathbf{S}, h\right) \ll 1$, then diffusion is dominating in a cell with diameter h. This is desirable, because then there is no need to introduce an extra diffusion term for stability.

It is important to notice that the Peclet Number is dependent on both the width of the cell, h, and the spatial values, or the underlying asset values, represented by $\mathbf{S}$. Since the value of $h \to 0$ as

the mesh is discretized more, it is sufficient to show that a related value, that I'll call the **Modified Peclet Number** is simply less than 1. I'll define the Modified Peclet Number, or $\mathbf{MPN}$ below as:

$$\mathbf{MPN}(\mathbf{S}) = \frac{\beta(\mathbf{S})}{\kappa(\mathbf{S})}. \tag{4.50}$$

Although the condition that the $\mathbf{MPN} < 1$ is not the same as the Peclet Number being much less than 1, it is more stringent and easier to analyze. Thus, I will use this more stringent condition to analyze the Peclet Number of the Basket-Case PDE.

Plotting the $\mathbf{MPN}$ in a 2-dimensional domain, I see two important features. One, is that the $\mathbf{MPN}$ is very large for values of S close to 0. The second feature I notice in this plot is that $\mathbf{MPN} \to 0$ as $\mathbf{S} \to \infty$. The first observation initially indicates that the PDE is advective dominated in this region, and thus may require some stabilization techniques. It turns out that, although the PDE is advection dominated in the region close to the origin, I do not need to use any stabilization techniques here. One reason for this is because the solution is so smooth here. The boundary conditions here are a constant value of 0. When looking at this through the lens of a Fourier Analysis, I can see that there are virtually no high frequency components to the solution in this area. This means that, although the PDE is advection dominated here, there are virtually no high frequency components to cause instability in the solution. Thus, no stabilization is needed here. The second observation indicates that very quickly after the values of the underlying assets begin to have value, the $\mathbf{MPN}$ falls below 1. This tells me that when the values of the underlying assets are not close to 0, then the PDE is diffusion dominated and thus needs no stabilization techniques.

Altogether, after analyzing the $\mathbf{MPN}$, I can see that I need no stabilization techniques to solve this PDE, and thus my FEM discretization should result in a stable solution.

4.6 Chapter Summary

In this chapter, I developed the weak formulation for the Basket-Case Black-Scholes Model.

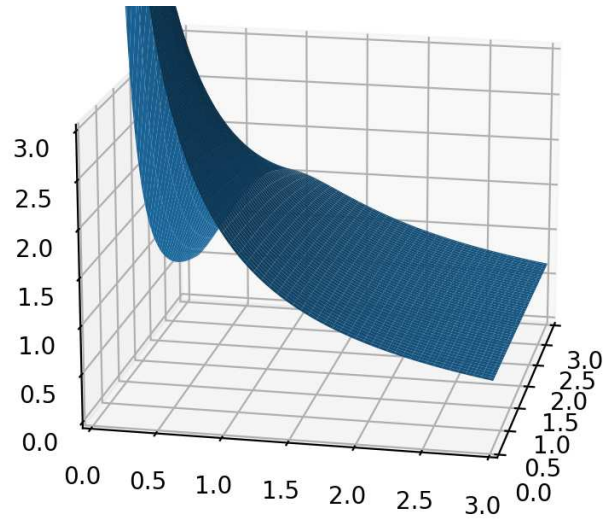


Figure 4.1: Modified Peclet Number 2-D.

I started by truncating the spatial domain of my model. This was necessary in order for me to apply the Finite Element Method. I simply chose a value that was "large-enough" in order for me to understand the behavior of the solution to the PDE.

Next, I reformulated the PDE in order to help with the construction of the weak formulation. Originally, I had the PDE stated in its non-conservative form. To make the weak formulation easier to construct, I decided to convert the PDE to its conservative form. This was because the conservative form was much easier to integrate-by-parts in order to get rid of the second derivative of the solution, V^n.

Then, I began the time-discretization. I began by partitioning the time domain into several equally sized intervals. Then, with this partition, I used an approximation for the time derivative of the value of my option. Finally, I used a *theta-method* to discretize the rest of the PDE with respect to time.

Next, I discretized my spatial domain. I did this by also partitioning the space domain into several equally sized "cells". Then, I constructed some piece-wise linear basis functions on each of these cells. With these basis functions, I was able to construct a weak formulation by multiplying my PDE by these basis functions and integrating over the domain.

Then, to complete my weak formulation, I had to choose values for the various thetas. These values depended on the characteristics of each piece of the weak formulation that the thetas corresponded to.

Finally, I discussed the reason for no stability terms being necessary in my numerical approach. The reason was because the Basket-Case PDE is diffusion-dominated. I used the Peclet Number to analyze the need for stability terms in my numerical approach, and decided that I did not need artificial stabilization because of the PDE being diffusion-dominated.

With this weak formulation, I can now use a computer program to solve for the various coefficients. For my thesis, I used a `C++` library called `deal.II` to write a program that solves for the coefficients using the weak formulation given in Equation (4.45). The results of this program will be discussed in the next chapter.

Chapter 5

Results

Taking the weak formulation that I generated in Section 4.4, I wrote a `C++` program using the Finite Element Library called `deal.II` [5]. The source code for this program can be found on my github page [2].

In this chapter, I will analyze the output from my program. First, I will look at two different views of the three-dimensional solution. Looking at a solution with three spatial dimensions and one time dimension would require a four-dimensional plot. I cannot hope to look at the entire solution then, so I will look at two different views of the solution that can still give me some insight into how the solution behaves.

Next, I will describe a way that I verified that the code was implemented in the way I described. I will also show some output from this process and I will describe it in detail.

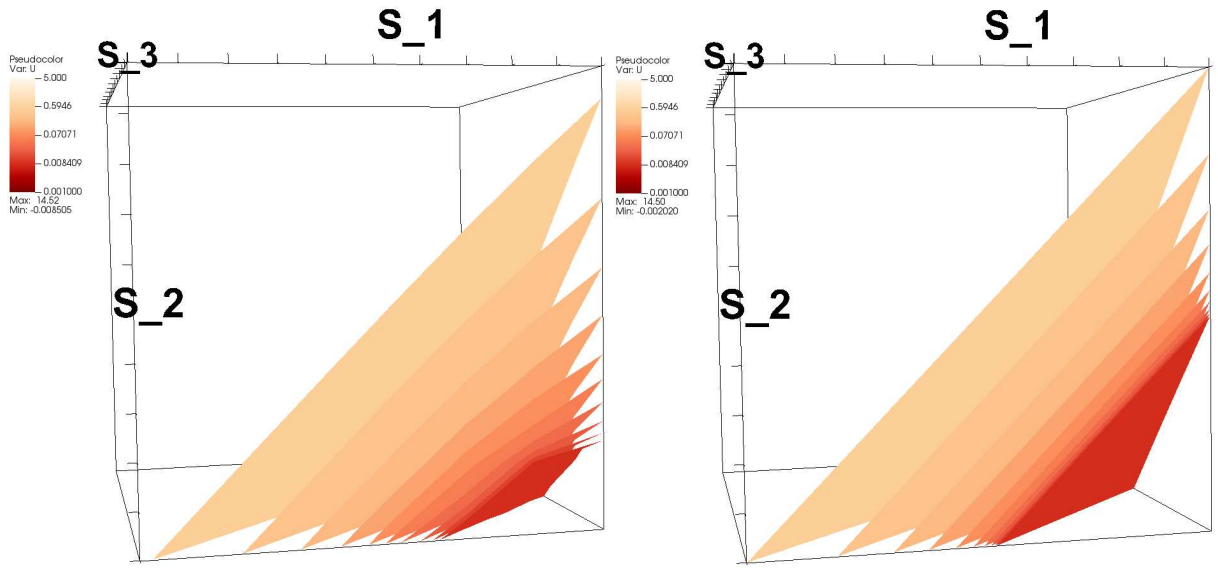
Finally, I will talk about the convergence of my numerical solution. In particular, I will be concerned with the H^1 and L^2 convergence. I will also discuss the numerical error some.

5.1 Three-Dimensional Solution

This section will focus on the visualization of the solution to the Basket-Case Black-Scholes Model. As I stated in the introduction to this chapter, I will have to be creative in order to visualize the solution. The interesting thing in the solution, in fact, is the "kink" that is present along the plane formed by

$$S_1 + S_2 + S_3 = K \tag{5.1}$$

at the initial time. Therefore, I will present two visualizations to show this "kink" and to explain the physical interpretation of it. The first image I will present is a look at the iso-surfaces with solution values that are small. By looking at these iso-surfaces, I hope to show how the kink gets smoothed out.



(a) Initial Time. **(b)** Expiration Time.

Figure 5.1: Iso-Surfaces of solution to 3D Basket-Case Model.
Iso-surfaces of the solution to a three dimensional Basket-Case option with strike price $K = \$0.5$.

The second image I will present is an elevation of a two dimensional slice of the domain. I will choose a non-boundary slice and show that the kink gets smoothed out in this case as well. With this elevated slice, I will also be able to visualize a part of the kink very clearly.

The reason that I am concerned with the kink is because this is the most interesting part of the solution. There are several reasons for this, but perhaps a simple explanation for this is the following. Recall from Section 4.5, the PDE behaves in a largely diffusive manner. Therefore, I should expect that the initial condition with the kink should smooth out some as time advances. This is not a rigorous proof, but for the purposes of my thesis, this intuition is an explanation for why I am interested in this kink.

I will begin by looking at the iso-surfaces of the solution.

5.1.1 Iso-Surfaces

The first images I will look at are found in Figure 5.1. These are images of iso-surfaces of the solution when the solution values are between zero and five. I chose these values because I wanted

to look at iso-surfaces for low values of the option, which is where I will be able to see properties of the kink in the solution.

In these images, I see that at the expiration time, the iso-surfaces are all parallel planes. In contrast, at the initial time, these iso-surfaces are not the same. At the initial time, these iso-surfaces start to curve towards the origin and they curve more as the option's value is less.

This shows the smoothing out of the kink in the solution. At the expiration time, the value of the option is either zero, or it is a linear function. This means that the iso-surfaces at the expiration time should all be flat planes. The option will suddenly have a non-zero value when the sum of the values of the underlying assets is greater than the strike price. This is seen in Figure 5.1b because the iso-surfaces are all flat planes, including the iso-surface for when the solution has a value of zero.

In contrast, the iso-surfaces at the initial time look much different than they do at the expiration time. These surfaces have a curvature to them, and get more curved as the solution value gets closer to zero. From a diffusive perspective, this shows the smoothing out of the kink that was present at the expiration time. This kink becomes a curve that becomes more curved as the sum of the asset values approaches the strike price.

Physically, this has a very nice interpretation. At the expiration time, the value of the option is very clear. Given the values of the underlying assets, I can give an exact value for the option because the values of the underlying assets are known. As I get further away from the expiration time, however, the value of the option becomes less clear because the values of the assets could change. I do not know what they will be worth at the expiration time.

If the value of the option is very large, then the values of the underlying assets must be very large compared to the strike price. In this case, the random fluctuations in the values of the assets doesn't matter very much, because they are already very valuable and it is unlikely that they will fall dramatically in value. This is due to the distribution of the values of each asset being a log-normal distribution. This is seen in the iso-surfaces being relatively flat as the value of the option gets further from zero. Because the fluctuations in the asset values doesn't matter very much in

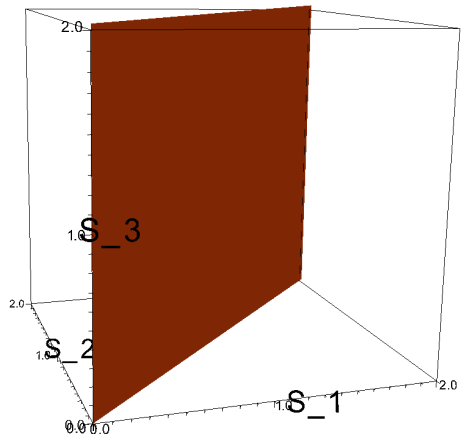


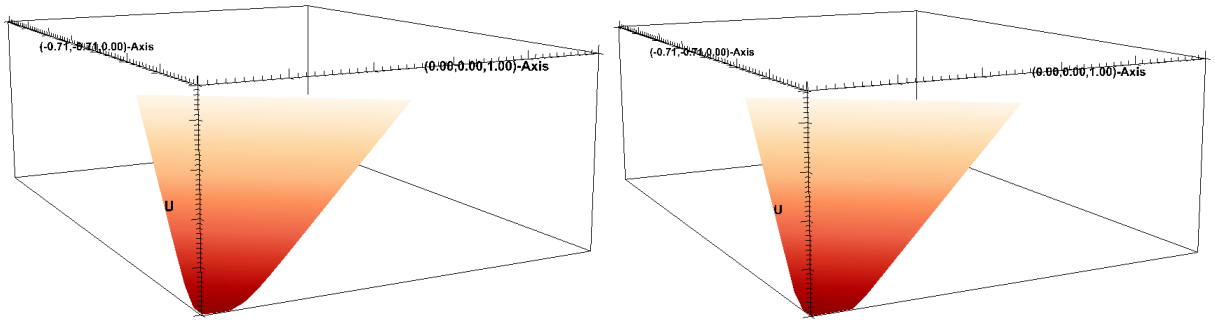
Figure 5.2: Two-Dimensional Slice of Domain.

this case, then the iso-surfaces will not be distorted very much from what they looked like at the expiration time.

If, on the other hand, the value of the option is very low, then the values of the underlying assets sum up to be very close to the strike price. In this case, the fluctuations in the asset values could cause them to sum up to something other than the strike price. This would cause the value of the option to be either worthless, or worth something. Unfortunately, it is not so clear what the value of the option will be in this case. So, the randomness in the asset values heavily distorts the iso-surfaces and results in introducing a curvature to them.

In conclusion, the iso-surfaces show that the kink was smoothed out from the expiration time. This "smoothing out" effect comes from the diffusive nature of the Basket-Case Black-Scholes PDE. At expiration time, the iso-surfaces are all flat planes, but become curved at the initial time. This curvature, becomes more pronounced as the value of the option approaches zero as well.

In the next section, I will look at one more image, which is the elevated slice. This image will actually show a kink in the solution that gets smoothed out.



(a) Initial Time for Elevated Slice. **(b)** Expiration Time for Elevated Slice.

Figure 5.3: Elevated Slice at Initial Time and Expiration Time.
Two dimensional elevated slices with $K = \$0.5$.

5.1.2 Two-Dimensional Slice

In this section, I will look at another visualization of the solution. I will take a two-dimensional slice of the domain, and then plot the solution in this sliced domain. I have chosen to take a slice that is perpendicular to the $S_1 - S_2$ plane and is at an angle of $45°$ off of the S_1 axis. This slice is visualized in Figure 5.2. I chose this slice because it will cut through the kink in the solution, and therefore will give me a way to visualize part of the kink.

The elevated solution is plotted in Figure 5.3, with the initial and expiration times of the solution shown. At the expiration time, I can clearly see the kink in the solution where the value of the option goes from zero to something non-zero. Then, at the initial time, I see that this kink has been smoothed out. There is now a gradual curve going from a worthless option to a non-zero valued option.

Plotting the solution along a slice of the domain allows for me to visualize the kink in the solution as well as the smoothing out of the kink.

5.2 Method of Manufactured Solutions

In this section, I will discuss how my code was verified. Because I do not have a three-dimensional solution to verify against, I must find another way to provide confidence that my code is correct.

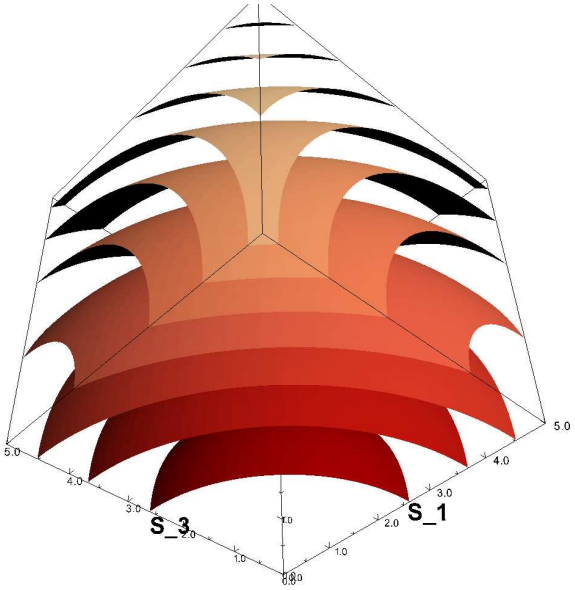


Figure 5.4: MMS Solution Iso-Surfaces.
Iso-surfaces for $f(S_1, S_2, S_3, \tau) = 6 - S_1^2 - S_2^2 - S_3^2 - \tau^2$.

I will use something called the *Method of Manufactured Solutions* to verify not only the weak formulation that I created in Section 4.4, but also the order of accuracy of my numerical solution [23].

This approach, in essence, requires me to *choose* an analytical solution. This choice of solution will bring with it boundary conditions and forcing terms. Then, I run my code against this new problem to verify that my program constructs the chosen solution. If it does, then I can have confidence that my code was implemented correctly. In addition to this, I can also refine the mesh in a systematic way to calculate the convergence rate of the numerical solution.

The steps for how to perform this have been done many times, and I will direct the reader to a paper by Patrick J. Roache for the steps and an example for how this process is done [23].

For my thesis, I will choose the analytic solution of

$$V(\mathbf{S}, \tau) = 6 - \sum_{i=0}^{3} S_i^2 - \tau. \tag{5.2}$$

This solution is a three-dimensional version of an upside down parabola. This solution was chosen because it is non-trivial, but also has an easily recognizable solution.

In Figure 5.4, I have plotted the iso-surfaces of the solution generated by my program using the Method of Manufactured Solutions. Here, I can see that the iso-surfaces are that of a three-dimensional upside down parabola. This gives me some amount of confidence, at least visually, that the code is correct.

The importance of this method is that it gives me a way to verify that my code is implemented correctly. I selected a non-trivial analytical solution to use in this method. Then, I visually verified that the solution was correct by looking at the iso-surfaces of the numerical solution generated. Next, I will use the Method of Manufactured Solutions to calculate the convergence of my numerical scheme and get the errors in the numerical solution.

5.2.1 Convergence and Error

It is good to visually inspect that the code is implemented correctly, but I also need to make sure that my numerical scheme will converge to the correct solution. By this, I mean that as I refine the mesh, I should get closer and closer to the exact solution to the Basket-Case Black-Scholes Model.

To do this, I used the Method of Manufactured Solutions again, but this time I solved on several different meshes. In each iteration, I performed a global mesh refinement and then ran my numerical simulation again and calculated the error. The results from this process are seen in Tables 5.1 and 5.2.

In Table 5.1, I can see that the global error with respect to the L^2 and L^∞ norms get smaller as the mesh is refined. I also see this same effect when looking at the H^1 semi-norm. From an application perspective, this is desirable because it shows that I should be able to accurately calculate the value of a three-asset Basket-Case Model if I simply refine my mesh enough. It should be noted that this "good enough" accuracy may be very computationally expensive, however.

In Table 5.2, I see that the convergence rates with respect to the L^2 norm and H^1 semi-norm are what is to be expected based on common numerical analysis theory [9]. Two important columns that I will point out are the Error Reduction Rate (ERR) column and the $\log_2$(ERR) column.

Table 5.1: Error Table.
Table showing relationship between the refinement of the mesh and the global errors with respect to various norms and semi-norms.

# cells	# dofs	L^2-error	H^1-error	L^∞-error
1	8	1.314×10^2	5.649×10^1	1.643×10^1
8	27	3.110×10^1	2.821×10^1	4.118
64	125	7.615	1.402×10^1	1.036
512	729	1.878	6.995	2.598×10^{-1}
4096	4913	4.654×10^{-1}	3.495	6.505×10^{-2}
32768	35937	1.146×10^{-1}	1.747	1.627×10^{-2}

The ERR column describes how the error has reduced with respect to the previous error. With respect to the H^1 semi-norm, I see that the error gets cut in half for each global refinement that is done on the mesh. With respect to the L^2 norm, the error seems to drop by a factor of four with each global refinement that is done on the mesh.

The $\log_2$(ERR) column is important to look at because of what a global refinement means. With each global refinement done on the mesh, the diameter of each cell, given by h, is made smaller by a factor of 2. Therefore, I can use this column to see what the convergence rate with respect to different norms and semi-norms is. With respect to the H^1 semi-norm, I see that the convergence rate must be $\mathcal{O}(h)$. With respect to the L^2 norm, the convergence rate is $\mathcal{O}(h^2)$. These show me that, in addition to the error getting smaller as the mesh is refined, I also have some sense as to how *much* smaller the error gets with each refinement. This gives me an idea for how much refinement I will need if I need a certain level of accuracy.

This section has covered how my numerical solution converges and what the errors are at various levels of mesh refinements. I see that the numerical solution proposed in my thesis converges on the order of $\mathcal{O}(h^2)$ with respect to the L^2 norm, and converges on the order of $\mathcal{O}(h)$ with respect to the H^1 semi-norm. These are exactly what should be expected based on numerical analysis theory.

Table 5.2: Convergence Rate Table.
Table showing the convergence rates with respect to the H^1 semi-norm and the L^2 norm.

	H^1-error			L^2-error		
# cells	Error	ERR	$\log_2$(ERR)	Error	ERR	$\log_2$(ERR)
1	5.649×10^1	-	-	1.314×10^2	-	-
8	2.821×10^1	2.00	1.00	3.110×10^1	4.22	2.08
64	1.402×10^1	2.01	1.01	7.615	4.08	2.03
512	6.995	2.00	1.00	1.878	4.05	2.02
4096	3.495	2.00	1.00	4.654×10^{-1}	4.04	2.01
32768	1.747	2.00	1.00	1.146×10^{-1}	4.06	2.02

Chapter 6

Summary and Conclusion

In this thesis I have discussed a numerical solution to the Basket-Case Black-Scholes Model with three underlying assets. In particular, this thesis focused on pricing a European Call Basket-Case Option.

I began my thesis by giving an introduction into basic Stochastic Calculus that was necessary in deriving the model in Chapter 2. This chapter gave many theorems that ultimately resulted in an understanding of Itô's Lemma, which was a crucial lemma in deriving the model.

Then, in Chapter 3, I gave a derivation of both the one and multi-dimensional model. The one-dimensional model, which was originally proposed in [8], was important for understanding the derivation of the multi-asset model. In this chapter, I also gave discussions on the rather interesting boundary conditions that were present in the Basket-Case version of the model. These boundary conditions were interesting because they were lower-dimensional solutions. This resulted in having to solve not one, but seven Black-Scholes models in order to get a numerical solution to the three-dimensional model. This chapter ended with a complete statement of the three-dimensional Basket-Case Model.

Next, in Chapter 4, I presented a numerical scheme that I employed to construct a numerical solution to the three-dimensional Basket-Case Model. I used a theta-method to discretize the time domain, and then I used an IMEX scheme to discretize the spatial domain. An IMEX scheme was used in order to handle the diffusion and advection terms separately. I also discussed the Peclet Number for this model, which gave a reasoning for no stabilization terms being necessary in my numerical scheme. This chapter concluded with a fully-discrete formulation to the Basket-Case Model.

Finally, in Chapter 5, I presented the results from using the fully discrete formulation I presented in Chapter 4. I presented a few figures to show the 'kink' in the solution. Then, I discussed how I used the Method of Manufactured Solutions to validate my code. I used an upside down

parabola as my manufactured solution, and I saw that my program was able to construct a numerical solution for this manufactured solution correctly. Finally, I used this method to also calculate errors and convergence rates. I saw that my numerical scheme was able to achieve a convergence rate on the order of $\mathcal{O}(h^2)$ with respect to the L^2 norm, and a convergence rate on the order of $\mathcal{O}(h$ with respect to the H^1 semi-norm. These convergence rates are what is expected based on numerical analysis theory.

In conclusion, the Finite Element Method is a useful tool in constructing a numerical solution to the Basket-Case Black-Scholes Model with three underlying assets. Although other numerical methods may be used to construct a solution to this model, I chose the FEM because of its use in solving other diffusion-dominated PDEs and because of my interest in this method. This method does become difficult to use when the option covers more than three underlying assets, however it may be useful for options with a low number of underlying assets. Employing this method has given me a greater understanding of the PDE in this model, and how to construct numerical solutions to PDEs in practice using the FEM.

References

[1] Kaushik I Amin and Robert A Jarrow. "Pricing options on risky assets in a stochastic interest rate economy 1". In: *Mathematical Finance* 2.4 (1992), pp. 217–237.

[2] Tyler Anderson. "tyand96/Black-Scholes-Multi-Dim-General: Thesis Release". In: (2023). DOI: 10.5281/zenodo.7542620.

[3] R Anguelov, P Kama, and JM-S Lubuma. "On non-standard finite difference models of reaction–diffusion equations". In: *Journal of computational and Applied Mathematics* 175.1 (2005), pp. 11–29.

[4] Julia Ankudinova and Matthias Ehrhardt. "On the numerical solution of nonlinear Black–Scholes equations". In: *Computers & Mathematics with Applications* 56.3 (2008), pp. 799–812.

[5] Daniel Arndt et al. "The `deal.II` Library, Version 9.4". In: *Journal of Numerical Mathematics* 30.3 (2022), pp. 231–246. DOI: 10.1515/jnma-2022-0054. URL: https://dealii.org/deal94-preprint.pdf.

[6] Alan H Barr. "The Einstein summation notation". In: *An Introduction to Physically Based Modeling (Course Notes 19), pages E* 1 (1991), p. 57.

[7] Hamid Naghibi Beidokhti et al. "A comparison between dynamic implicit and explicit finite element simulations of the native knee joint". In: *Medical engineering & physics* 38.10 (2016), pp. 1123–1130.

[8] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *The Journal of Political Economy* 81.3 (1973), pp. 637–654.

[9] Susanne C Brenner, L Ridgway Scott, and L Ridgway Scott. *The mathematical theory of finite element methods*. Vol. 3. Springer, 2008.

[10] Hans-Joachim Bungartz et al. “Parallelizing a Black-Scholes solver based on finite elements and sparse grids”. In: *2010 IEEE International Symposium on Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW)*. IEEE. 2010, pp. 1–8.

[11] Jehanzeb H Chaudhry et al. “A posteriori error analysis of IMEX multi-step time integration methods for advection–diffusion–reaction equations”. In: *Computer Methods in Applied Mechanics and Engineering* 285 (2015), pp. 730–751.

[12] Les Clewlow and Stewart Hodges. “Optimal delta-hedging under transactions costs”. In: *Journal of Economic Dynamics and Control* 21.8-9 (1997), pp. 1353–1376.

[13] R. Deep. *Probability and Statistics: With Integrated Software Routines*. Elsevier Science, 2005, p. 108. ISBN: 9780123694638. URL: https://books.google.com/books?id=FCax7O6U6NQC.

[14] I Maruddani Di Asih et al. “Modeling stock prices in a portfolio using multidimensional geometric brownian motion”. In: *Journal of Physics: Conference Series*. Vol. 1025. 1. IOP Publishing. 2018, p. 012122.

[15] Lawrence Evans. *An introduction to stochastic differential equations*. Providence, Rhode Island: American Mathematical Society, 2013. ISBN: 9781470410544.

[16] Marijke Huysmans and Alain Dassargues. “Review of the use of Péclet numbers to determine the relative importance of advection and diffusion in low permeability environments”. In: *Hydrogeology Journal* 13.5 (2005), pp. 895–904.

[17] Kiyosi Itô. “109. stochastic integral”. In: *Proceedings of the Imperial Academy* 20.8 (1944), pp. 519–524.

[18] Humam Kareem Jalghaf et al. “Explicit Stable Finite Difference Methods for Diffusion-Reaction Type Equations”. In: *Mathematics* 9.24 (2021), p. 3308.

[19] Joonglee Jo and Yongsik Kim. “Comparison of numerical schemes on multi-dimensional Black-Scholes equations”. In: *Bulletin of the Korean mathematical society* 50.6 (2013), pp. 2035–2051.

[20] Riccardo Mannella and Peter VE McClintock. “Itô versus Stratonovich: 30 years later”. In: *Fluctuation and Noise Letters* 11.01 (2012), p. 1240010.

[21] Robert Mullen and Ted Belytschko. “Dispersion analysis of finite element semidiscretizations of the two-dimensional wave equation”. In: *International Journal for Numerical Methods in Engineering* 18.1 (1982), pp. 11–29.

[22] Wei Qu and Yong Liang. “Stability and convergence of the Crank-Nicolson scheme for a class of variable-coefficient tempered fractional diffusion equations”. In: *Advances in Difference Equations* 2017.1 (2017), pp. 1–11.

[23] Patrick J Roache. “Code verification by the method of manufactured solutions”. In: *J. Fluids Eng.* 124.1 (2002), pp. 4–10.

[24] Kambiz Salari and Patrick Knupp. *Code verification by the method of manufactured solutions*. Tech. rep. Sandia National Lab.(SNL-NM), Albuquerque, NM (United States); Sandia . . ., 2000.

[25] Y. Suhov and M. Kelbert. *Probability and Statistics by Example*. Probability and Statistics by Example v. 1. Cambridge University Press, 2014. ISBN: 9781107603585. URL: https://books.google.com/books?id=JftkBAAAQBAJ.

[26] Tamás Szabados. “An elementary introduction to the Wiener process and stochastic integrals”. In: *arXiv preprint arXiv:1008.1510* (2010).