Barrier Option Pricing Under Meromorphic Lévy Processes Observed at Poisson Arrival Times

Jesse Benjamin Tonkin

April 2023

Thesis submitted for the degree of
Master of Philosophy
in
Applied Mathematics
at The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences
School of Mathematical Sciences

THE UNIVERSITY
of ADELAIDE

Contents

Signed Statement **ix**

Acknowledgements **xi**

Abstract **xiii**

1 Introduction **1**

2 Mathematical foundations **3**
2.1 Elementary measure theory 3
2.2 Stochastic processes 10
2.3 Lévy processes 11
2.3.1 Variation and jump activity 15
2.3.2 Examples 17
2.3.3 Lévy-Itô decomposition 22
2.3.4 Wiener-Hopf factorisation 23

3 Financial background **27**
3.1 Financial markets and securities 27
3.1.1 Derivatives 29
3.1.2 Efficient and complete markets 31
3.1.3 Barrier options 32
3.1.4 Bonds 33
3.2 Stochastic modelling of options 33
3.2.1 Discounted price process 34
3.2.2 Equivalent martingale measure 34
3.3 Black-Scholes formula 35

4 Option pricing models driven by Lévy processes **41**
4.1 Types of Lévy processes used in finance 42
4.1.1 KoBoL processes 42
4.2 Equivalent martingale measures in Lévy models 43

4.2.1 Esscher transform . . . 44
4.3 Option pricing models . . . 44
4.3.1 Laplace transform method . . . 46
4.3.2 Carr's randomisation . . . 53
4.4 Simple Wiener-Hopf method . . . 56
4.4.1 SWH method implementation . . . 61
4.5 Wiener-Hopf Monte Carlo simulation of Lévy processes . . . 65

5 Meromorphic Lévy processes 69
5.1 Meromorphic Lévy processes . . . 69
5.1.1 Characteristic exponent . . . 73
5.1.2 Wiener-Hopf factors . . . 73
5.1.3 Variance and jump activity . . . 75
5.2 Examples . . . 78
5.2.1 β-family processes . . . 78
5.2.2 Hyper-exponential processes . . . 82
5.2.3 Other examples . . . 83

6 Meromorphic Wiener-Hopf method 85
6.1 Numerical implementation of meromorphic processes . . . 85
6.2 The Algorithm . . . 86
6.3 Simple Wiener-Hopf method under meromorphic processes . . . 92
6.4 Meromorphic Wiener-Hopf Monte Carlo simulation . . . 94
6.5 Evaluation of the MWH method . . . 99

7 Numerical experiments 101

8 Conclusion 105

A Appendix 109
A.1 Meromorphic function . . . 109
A.2 Fast Fourier transform . . . 110
A.3 Interval bisection method . . . 112
A.4 Acceptance-rejection sampling . . . 113

Bibliography 115

List of Tables

4.1 SWH results for a knock-out barrier put option with $K = 100$, $H = 90$, $r = 0.07231$, and $T = 0.5$, driven by a KoBoL process with $\mu = 0$, $\nu_+ = \nu_- = 0.5$, $c_+ = c_- = 1$, $\lambda_+ = 9$, and $\lambda_- = -8$. 65

7.1 Estimated option values using Monte Carlo simulations (with 95% confidence interval), the MWH method, and the SWH method. 103
7.2 Estimated option values using the MWH and SWH method. 103
7.3 Estimated option values using the MWH method with a non-Gaussian component. 104

List of Figures

2.1 Brownian motion realisations with $\mu = 0$ and $\mu = 1.4$, both with $\sigma^2 = 4$. . 17
2.2 Poisson process realisation with rate $\lambda = 5$. 19

3.1 TSLA stock price evolution over the 12-month period in 2021. 28
3.2 Examples of potential put option contracts. Green and red indicate periods with positive and zero payout, respectively. 30
3.3 Examples of potential barrier put option contracts. Green and red indicate periods with positive and zero payout, respectively. 33

4.1 Barrier option depicting regions B and C in red and green, respectively. . . 45

6.1 Plot of roots of $q + \Psi(i\xi) = 0$ for a β-family Lévy process with $q = 1$. . . . 91
6.2 Plot of subordinators, X^+ and X^- that decompose a meromorphic Lévy process. 94
6.3 Probability density of $S_j = \overline{X}_{\mathbf{e}_1 T/M}$, $f_S(x)$, and a dominating function $g_S^*(x)$ over $x \in [0, 10]$ for a β-family Lévy process with $\mu = 1$, $\sigma = 0.5$, $\alpha_1 = \alpha_2 = 1$, $\beta_1 = \beta_2 = 1.5$, $\lambda_1 = \lambda_2 = 1.5$, $c_1 = c_2 = 1$ and $N = \widehat{N} = 1000$. The observation rate for the roots is $q = M/T = 500/2$. 97
6.4 Plot of five independent realisations of a β-family Lévy process using the WHMC techniques for meromorphic Lévy processes. 99

7.1 Option price calculated for the parameters given in Tables 7.1 and 7.2 using the MWH method with varying number of roots, $N \in [50, 500]$. 104

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the The University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed: ..... Date: 05/04/2023.....

Acknowledgements

I would first like to express my deepest appreciation to my supervisors, Giang Nguyen and Oscar Peralta, who both provided me with unwavering guidance and support. From day one, you cultivated a welcoming environment where I felt confident and included. I have developed tremendously thanks to you, both as a mathematician and as a person.

I am thankful to my entire family, particularly my parents, Joy and Kelvin Tonkin, who have always supported me. You regularly showed interest in my research, even if there was no hope of understanding it. My dog, Shep, also deserves a mention; he was always there for me, especially around dinner time.

Finally, there were numerous close friends who helped keep me sane through my studies. I am grateful to Wills Nguyen for his continual encouragement to actually attend social events, my university experience was certainly improved as a result. Thanks also to Matt Innes and Ben Norris for, perhaps too often, distracting me with games when I needed a break from research.

For the many others whom I could not mention by name, I am abundantly appreciative of the impact that you have on me and look forward to my future endeavours with all of you.

Abstract

Calculating the expected value of *options* is a prominent task in financial mathematics literature. *Exotic options*, such as *barrier options*, provide a further specialisation that often requires vastly differing techniques. The first explicit continuous-time option pricing formula, the *Black-Scholes model*, assumed that the underlying *security* could be modelled using a *stochastic process* called *Brownian motion*. While this was reasonably effective, empirical research has shown that financial markets have path properties and distributions that cannot be achieved by Brownian motion. Accordingly, over the past two decades, the literature has investigated the use of other *Lévy processes* in option pricing models. In these models, the expected value of barrier options may be represented through a system of *integro-differential equations*, which can be solved numerically using the *Wiener-Hopf factors* of the process. However, many Lévy processes do not have known expressions for the Wiener-Hopf factors, so previous algorithms, such as the *Simple Wiener-Hopf* (SWH) method [33], have required methods of approximating these functions. Additional calculations and errors in these approximations reduce the accuracy and time efficiency of these algorithms.

In this thesis, we investigate the family of *meromorphic Lévy processes*, which have known Wiener-Hopf factors and are more suitable than Brownian motion. We derive an algorithm to price barrier option when the model is driven by a meromorphic process, calling this new algorithm the *Meromorphic Wiener-Hopf* (MWH) method. By comparing the results with *Monte Carlo simulations* and the SWH method, we demonstrate that the MWH method provides an accurate estimate of barrier option prices. While there is currently no significant difference in computational speed between the MWH and SWH methods, we suggest potential means to optimise the MWH method. Regardless, the MWH method is a beneficial addition to the suite of option pricing models because it may be applied in different conditions than other algorithms. In particular, the MWH method allows for processes of either finite or infinite *path variation*, while the SWH requires finite variation.

Chapter 1

Introduction

Mathematical finance is an area of mathematics concerned with modelling financial markets. A prevalent problem in mathematical finance is the modelling of financial *assets* to estimate the value of *derivatives* such as *options*. There are many types of options, the most common are *American* and *European*. However, various other *exotic options* exist, such as *lookback*, *compound*, and *barrier options*, each with differing properties. This research is particularly interested in barrier options. The variety of option types leads to unique and interesting behaviours which we want to understand mathematically, often achieving this using *stochastic models*.

The first explicit model of continuous-time financial markets was the *Black-Scholes model*, which was based on a mathematical object called *Brownian motion*. Stochastic models of financial markets have evolved significantly since then; a key advancement over the past two decades has been the implementation of models driven by *Lévy processes* that are different from Brownian motion. The resultant models are improved because there are several unreasonable assumptions in the Brownian motion model that can be removed using Lévy processes.

There is a vast literature on methods for pricing options under Lévy processes, which can be categorised into three groups: Monte Carlo simulation (see, *e.g.*, [31, 37, 56]), backward induction (see, *e.g.*, [22, 27, 50]), and integro-differential equations (see, *e.g.*, [10, 32, 43]). The latter, where our research lies, may be solved using a *Fourier inversion* technique that depends on the *Wiener-Hopf factors* of the process. However, explicit forms of the Wiener-Hopf factors are not known for many Lévy processes. Several techniques have been proposed to approximate these functions, and hence derive an option pricing model (see, *e.g.*, [32, 33]). However, these approximations reduce the accuracy and speed of the algorithms, and often require assumptions that limit the flexibility of the model.

In recent years, a subclass of Lévy processes has been defined, called *meromorphic Lévy processes* [36]. Elements of this class have a known semi-explicit representation for their Wiener-Hopf factors in the form of an infinite product. The aim of this research

is to derive a technique to calculate the value of barrier options with a model driven by meromorphic Lévy processes, using the known form of the Wiener-Hopf factors. This would provide access to a new, extensive class of Lévy-driven financial models, which may allow for a better fit to security price dynamics. Additionally, since we have semi-explicit formulas, the algorithm may be faster and more accurate than other methods that approximate the Wiener-Hopf factors.

This thesis will present the relevant mathematics and finance to build up the problem of pricing options, then discuss meromorphic Lévy processes and investigate their properties, before finally deriving a barrier option pricing method. We first introduce the mathematical and financial background in chapters 2 and 3, respectively. Next, in Chapter 4, we discuss barrier option pricing models driven by Lévy processes, with reference to a recent algorithm from the literature. We then present meromorphic Lévy processes and prove several interesting properties in Chapter 5. Utilising the aforementioned techniques, we derive, in Chapter 6, an algorithm to calculate the value of a barrier option when modelled according to a meromorphic Lévy process. Finally, we demonstrate the new algorithm with numerical experiments in Chapter 7, comparing the results to other algorithms in the literature.

Chapter 2

Mathematical foundations

Measure theory is the study of mathematical measures, generalising the notions of length, area, and volume. Probability theory may be considered a branch of measure theory that is concerned with the probabilities of events, often essential in any field requiring quantitative analysis of data. Stochastic modelling is a key subject within probability theory that aims to understand and evaluate processes that evolve randomly. This has importance in a wide variety of applications, such as weather, biochemistry, and finance.

In this chapter, we provide the mathematical background that will allow us to build and justify our models. First, basic measure-theoretic ideas will be introduced, leading to definitions of the stochastic processes that are used in our model. This information is available in many textbooks on probability theory; see, *e.g.*, Bogachev [6], Billingsley [3], and Kallenberg [25]. Then, we discuss various theorems and mathematical properties that will be required. The majority of these statements are also common; our main resources are [38] and [51].

2.1 Elementary measure theory

Measure theory provides the setting for us to investigate random processes and it is important to understand its basic notations, definitions, and theorems. In this section, we define various measure-theoretic objects such as sets, measures, and σ-fields. We then discuss probability spaces, which are formed by these objects, along with random variables.

First, a *set* is a collection of elements, which can be mathematical objects such as numbers, symbols, or variables. A set with no elements is called the empty set, $\varnothing$. Other common sets are the real numbers, $\mathbb{R} = (-\infty, \infty)$, the natural numbers, $\mathbb{N} = \{1, 2, \dots\}$, and the integers, $\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$. We use $\mathbb{R}_+ = [0, \infty)$ to denote the non-negative real numbers, $\overline{\mathbb{R}} = [-\infty, \infty]$, the extended real line, and $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$, the non-negative integers. In a random experiment, the set of all possible outcomes, Ω, is called

the *sample space*. For example, in an experiment testing the number of heads in $n \in \mathbb{N}$ coin tosses, the result can be any integer number from 0 to n, so the sample space is $\Omega = \{0, 1, \ldots, n\}$. An element of Ω is called a *sample point*, and a subset of Ω is an *event*.

Given a sample space Ω, only some events $A \subseteq \Omega$ may be interesting, so we only want to assign probabilities to collections of these particular events. Denote the set of these collections by $\mathcal{F}$. Each element of $\mathcal{F}$, $A \in \mathcal{F}$, is a subset of the sample space, $A \subseteq \Omega$, meaning that $\mathcal{F}$ is a subset of the *power set* of Ω, $\mathcal{P}(\Omega)$. This set, $\mathcal{F}$, should satisfy the properties in Definition 1 (see [3, Section 2] for intuition behind these properties) and is called a *σ-field* (σ-algebra).

Definition 1 (σ-field)**.** *Let Ω be a set. Then, $\mathcal{F} \in \mathcal{P}(\Omega)$ is a **σ-field** if:*

i. $\Omega \in \mathcal{F}$.

ii. $\mathcal{F}$ *is closed under complements: if* $A \in \mathcal{F}$, *then* $A^c \in \mathcal{F}$, *where* $A^c = \Omega \backslash A$ *is the complement of the event* A.

iii. $\mathcal{F}$ *is closed under countable unions: if* $A_1, A_2, \ldots \in \mathcal{F}$, *then* $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$.

Together, these properties also imply that the empty set is in $\mathcal{F}$, $\varnothing \in \mathcal{F}$, and that $\mathcal{F}$ is closed under countable intersections.

Elements of a σ-field are called *measurable sets* and the ordered pair $(\Omega, \mathcal{F})$ is called a *measurable space*. On any set Ω, the smallest possible σ-field is $\{\varnothing, \Omega\}$ and the largest is $\mathcal{P}(\Omega)$. Many mathematical objects can be used to generate corresponding σ-fields with desirable properties. This is demonstrated below for sets and functions, and a third example is shown in Definition 14.

Definition 2. *Let Ω be a set and F be a collection of subsets of Ω. The **σ-field generated by the set** F, $\sigma(F)$, is the smallest σ-field containing every set in F, or equivalently, the intersection of all σ-fields that contain F.*

Definition 3. *Let X and Y be sets and T be a σ-field on Y. Define a function $f\colon X \to Y$. The **σ-field generated by the function** f, $\sigma(f)$, is*

$$\sigma(f) = \{f^{-1}(B)\colon B \in T\},$$

the set of all inverse images of f on B.

When the sample space is a topological space, we often use the *Borel σ-field*, which we define next.

Definition 4 (Borel σ-field)**.** *Given a topological space, E, the **Borel σ-field**, $\mathcal{B}(E)$, on E is the σ-field that is generated by the collection of open sets of E (or, equivalently, the closed sets).*

We can now define the notion of a *measure*, which generalises geometrical measures such as length, area, and volume, and is the basis of measure theory.

Definition 5 (Measure)**.** *Let E be a set and $\mathcal{E}$ be a σ-field on E. A function $\mu\colon \mathcal{E} \to \overline{\mathbb{R}}$ is a* ***measure*** *if:*

i. ***(Non-negativity)*** *For all $A \in \mathcal{E}$, $\mu(A) \geq 0$.*

ii. ***(Null empty set)*** $\mu(\varnothing) = 0$.

iii. ***(σ-additivity)*** *If $A_1, A_2, \ldots \in \mathcal{E}$ are pairwise disjoint, then*

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

A measurable space, $(E, \mathcal{E})$, along with a measure on that space, μ, forms a triplet, $(E, \mathcal{E}, \mu)$, which is called a *measure space*. Within probability theory, we have a further distinction of a measure, called a *probability measure*. This is simply a measure whose range is in $[0, 1]$ and where the image of the entire set is 1.

Definition 6 (Probability measure)**.** *Let Ω be a set and $\mathcal{F}$ be a σ-field on Ω. A function $\mathbb{P}\colon \Omega \to [0, 1]$ is a* ***probability measure*** *if:*

i. $\mathbb{P}$ *is a measure on* Ω.

ii. $\mathbb{P}(\Omega) = 1$.

These properties also imply that $\mathbb{P}(\varnothing) = 0$.

Given one probability measure, we often want to discuss its *equivalent measures*.

Definition 7 (Equivalent probability measures)**.** *Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ are* ***equivalent*** *if they have the same null sets. In other words, for all $A \in \mathcal{F}$,*

$$\mathbb{Q}(A) = 0 \quad \iff \quad \mathbb{P}(A) = 0.$$

We now have the mathematical building blocks to form a *probability space*, which is used to formally represent a random experiment.

Definition 8 (Probability space)**.** *A* ***probability space*** *is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a sample space, $\mathcal{F}$ is a σ-field on Ω, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.*

A singular outcome, ω, of a random experiment is an element of the sample space, Ω. An event $A \in \mathcal{F}$ is said to have occurred if it contains the outcome ω. The proportion of random experiments where A occurs is given by its probability, $\mathbb{P}(A)$. We are often interested in events that occur with probability one, or equivalently, occur *almost surely*. Since we can define non-empty subsets of the sample space to have probability zero, this does not imply that all possible outcomes belong to such an event.

Definition 9. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An event $A \in \mathcal{F}$ occurs $\mathbb{P}$-**almost surely** if $\mathbb{P}(A) = 1$, or equivalently, if $\mathbb{P}(A^c) = 0$. More generally, any event $E \subseteq \Omega$ occurs $\mathbb{P}$-almost surely if $E^c = \Omega \backslash E \subseteq N$ for some null set N (i.e. $N \in \mathcal{F}$ such that $\mathbb{P}(N) = 0$). Often, we omit the specification of the measure $\mathbb{P}$ and use an abbreviation, writing: A a.s..*

As a simple illustration of events that occur almost surely, consider a random experiment where we toss a (fair) coin repeatedly (an infinite number of times) and record each outcome as 'head' (H) or 'tail' (T). The probability space for each toss is $(\{H, T\}, 2^{\{H,T\}}, \mathbb{P})$, where $\mathbb{P}(H) = \mathbb{P}(T) = 1/2$. Letting the trials be independent and identically distributed (i.i.d.) and denoted by ω_1, ω_2, ..., an outcome of our experiment is a sequence of heads and tails, $(\omega_1, \omega_2, \ldots)$; any infinite sequence of heads and tails is a possible outcome. By independence, the probability measure of this experiment is $\mathbb{P}(\omega_1, \omega_2, \ldots) = \mathbb{P}(\omega_1)\mathbb{P}(\omega_2)\ldots$. Consider the event $A \in \mathcal{F}$ where we toss at least one head, such that A^c is the event where all tosses are tails. Then,

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \prod_{j \in \mathbb{N}} \mathbb{P}(T) = 1 - \prod_{j \in \mathbb{N}} \frac{1}{2} = 1.$$

Thus, by Definition 9, A occurs $\mathbb{P}$-almost surely. However, observe that the outcome of all tails, $(T, T, T, \ldots)$, does not belong to A. This shows that an event occurring almost surely does not imply that all possible outcomes belong to that event.

Since almost sure events are defined in terms of null sets, we often want to deal with probability spaces that have well-defined subsets of null sets, which are called *complete probability spaces*.

Definition 10 (Complete probability space)**.** *A probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, is **complete** if every subset of a set with measure zero is measurable, also with measure zero. In other words, for every $A \in \mathcal{F}$, if $\mathbb{P}(A) = 0$, then every subset of $S \subseteq A$ is such that $S \in \mathcal{F}$ and $\mathbb{P}(S) = 0$.*

Given a probability space, a *filtration* is another important mathematical object, which is used to represent the amount of available information at a given point. This will later contribute to the formalisation of *stochastic processes*, the basis of our models.

Definition 11 (Filtration)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and T be an index set with a total order, usually some subset of $\mathbb{R}_+$. For each $t \in T$, let $\mathcal{F}_t \subseteq \mathcal{F}$ be a σ-field, such that for each $k \leq \ell$, $\mathcal{F}_k \subseteq \mathcal{F}_\ell$. Then, $(\mathcal{F}_t)_{t \in T}$ is a **filtration**. Often we will omit the index set for a filtration and write $(\mathcal{F}_t)$.*

We can also define functions between measurable spaces such that the measurability structure of the spaces is preserved. This is analogous to the idea that continuous functions preserve topological structure by ensuring that the preimage of all open sets is still open. Similarly, a *measurable function* requires the preimage of measurable sets to still be measurable.

Definition 12 (Measurable function)**.** *Let (X, Σ) and $(Y, \mathcal{T})$ be measurable spaces. A function $f\colon X \to Y$ is* ***measurable*** *if, for every $B \in \mathcal{T}$*

$$f^{-1}(B) \coloneqq \{x \in X \mid f(x) = B\} \in \Sigma.$$

Finally, we are prepared to introduce the last basic object that is essential to probability theory: random variables. Often when performing a random experiment, we want to transform the sample point, usually to have a value given by a real number. This is achieved by defining random variables, which are functions from the sample space to some other measurable space. Any interesting subsets of outputs from our random variable should have defined probabilities (*i.e.*, should be measurable). Considering this, random variables are defined as follows.

Definition 13 (Random variable)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ be a measurable space. A measurable function $X\colon \Omega \to E$ is an* $(E, \mathcal{E})$***-valued random variable****. Equivalently, a function $X\colon \Omega \to E$ is an $(E, \mathcal{E})$-valued random variable if, for every $B \in E$, $X^{-1}(B) \in \mathcal{F}$.*

When the codomain of a random variable, E, is a topological space, the corresponding σ-field, $\mathcal{E}$, is usually the Borel σ-field. In fact, the codomain of the random variables in probability theory is often the real numbers, $\mathbb{R}$, which is a topological space, so we often use the Borel σ-field on $\mathbb{R}$, denoted $\mathcal{B}(\mathbb{R})$. In this setting, there is a useful alternative definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be the real numbers with the Borel σ-field. Then, a function $X\colon \Omega \to \mathbb{R}$ is a $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued random variable if

$$\{\omega \mid X(\omega) \leq r\} \in \mathcal{F} \quad \forall\, r \in \mathbb{R}.$$

This follows from Definition 13, since the set $\{(-\infty, r]\colon r \in \mathbb{R}\}$ generates the Borel σ-field. For brevity, we will subsequently refer to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-valued random variables without specifying the target space, simply saying *random variable*.

As with sets and functions previously, we can generate a σ-field from a random variable.

Definition 14. *Let X be an $(E, \mathcal{E})$-valued random variable, defined on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and measure space $(E, \mathcal{E})$. The* ***σ-field generated by the random variable****, X, is*

$$\sigma(X) = \left\{X^{-1}(B) \mid B \in \mathcal{E}\right\}.$$

Lastly, we define the *law* and *characteristic function* of a random variable. The law of a random variable is synonymous with its *probability distribution*.

Definition 15. *Let X be a random variable from the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The* ***law*** *of X is a probability measure, L_X, on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that, for all $B \in \mathcal{B}(\mathbb{R})$,*

$$L_X(B) = \mathbb{P}(X \in B).$$

It turns out, as we would hope, that the function defined to be the law of a random variable is always a probability measure. Proving this statement demonstrates several of our definitions.

Proposition 1. *The law of a random variable is always a probability measure.*

Proof. Let X be a random variable from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Define a function $L_X \colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ such that, for each $B \in \mathcal{B}(\mathbb{R})$,

$$L_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\}).$$

We need to show that L_X satisfies Definition 6, meaning:

i. L_X is a measure on Ω,

ii. $L_X(\mathbb{R}) = 1$.

Since X is a function with codomain $\mathbb{R}$, so that $L_X(\mathbb{R}) = \mathbb{P}(X \in \mathbb{R}) = 1$, condition (ii) is satisfied.

By Definition 5, for condition (i), L_X must have a null empty set and satisfy non-negativity and σ-additivity. The first two requirements are immediately satisfied because L_X is defined according to another probability measure, $\mathbb{P}$, which has these properties. To demonstrate σ-additivity, let $B_1, B_2, \ldots \in \mathcal{B}(\mathbb{R})$ be pairwise disjoint. Suppose that $\omega_1 \in \{\omega \in \Omega \mid X(\omega) \in B_1 \cup B_2 \cup \cdots\}$, so that $X(\omega_1) \in B_1 \cup B_2 \cup \cdots$. Then, there is at least one $i \in \mathbb{N}$ such that $X(\omega_1) \in B_i$. Hence, $\omega_1 \in \bigcup_{i \in \mathbb{N}} \{\omega \in \Omega \mid X(\omega) \in B_i\}$ and

$$\{\omega \in \Omega \mid X(\omega) \in B_1 \cup B_2 \cup \cdots\} \subseteq \bigcup_{i \in \mathbb{N}} \{\omega \in \Omega \mid X(\omega) \in B_i\}.$$

Now suppose that $\omega_2 \in \bigcup_{i \in \mathbb{N}} \{\omega \in \Omega \mid X(\omega) \in B_i\}$, so that there is at least one $i \in \mathbb{N}$ such that $X(\omega_2) \in B_i$. Then, $X(\omega_2) \in B_1 \cup B_2 \cup \cdots$. Hence, $\omega_2 \in \{\omega \in \Omega \mid X(\omega) \in B_1 \cup B_2 \cup \cdots\}$ and

$$\bigcup_{i \in \mathbb{N}} \{\omega \in \Omega \mid X(\omega) \in B_i\} \subseteq \{\omega \in \Omega \mid X(\omega) \in B_1 \cup B_2 \cup \cdots\}.$$

Thus,

$$\{\omega \in \Omega \mid X(\omega) \in B_1 \cup B_2 \cup \cdots\} = \bigcup_{i \in \mathbb{N}} \{\omega \in \Omega \mid X(\omega) \in B_i\}.$$

Additionally, since $B_1, B_2, \ldots$ are disjoint, $X(\omega) \in B_i$ for at most one $i \in \mathbb{N}$. Hence, for any $i \neq j$, the sets $\{\omega \in \Omega \mid X(\omega) \in B_i\}$ and $\{\omega \in \Omega \mid X(\omega) \in B_j\} \in \mathcal{F}$ are disjoint.

Then, by the σ-additivity of $\mathbb{P}$,

$$\begin{aligned} L_X\left(\bigcup_{i\in\mathbb{N}} B_i\right) &= \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B_1 \cup B_2 \cup \cdots\}) \\ &= \mathbb{P}\Big(\bigcup_{i\in\mathbb{N}} \{\omega \in \Omega \mid X(\omega) \in B_i\}\Big) \\ &= \sum_{i\in\mathbb{N}} \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B_i\}) \\ &= \sum_{i\in\mathbb{N}} L_X(B_i). \end{aligned}$$

So, L_X satisfies σ-additivity. Hence, L_X is a probability measure. □

Definition 16 (Characteristic function)**.** *Let X be a $\mathbb{R}$-valued random variable. The* ***characteristic function*** *of X is a function, $\theta\colon \mathbb{R} \to \mathbb{C}$, such that*

$$\theta(\xi) = \mathbb{E}\big[e^{i\xi X}\big] \quad \forall\, \xi \in \mathbb{R}.$$

The characteristic function uniquely identifies a random variable, which leads to Lemma 1. This result follows immediately from the *inversion formula* for the characteristic function that is discussed later, in Theorem 4.

Lemma 1. *Let X and Y be random variables with characteristic functions θ_X and θ_Y, respectively. Then X is equal in distribution to Y if and only if $\theta_X(\xi) = \theta_Y(\xi)$ for all $\xi \in \mathbb{R}$.*

Finally, Lemma 2 may be used to calculate the characteristic function of a sum of random variables. We will later use this result to apply Lemma 1.

Lemma 2. *Let $X_1, X_2, \ldots$ be independent random variables and $Y = X_1 + X_2 + \cdots$ with characteristic functions $\theta_{X_1}, \theta_{X_2}, \ldots$ and θ_Y, respectively. Then, for all $\xi \in \mathbb{R}$,*

$$\theta_Y(\xi) = \prod_{i\in\mathbb{N}} \theta_{X_i}(\xi).$$

Proof. Let $X_1, X_2, \ldots$ be independent random variables and $Y = X_1 + X_2 + \cdots$ with characteristic functions $\theta_{X_1}, \theta_{X_2}, \ldots$ and θ_Y, respectively. Then, by Definition 16 and the independence of X_i,

$$\theta_Y(\xi) = \mathbb{E}\big[e^{i\xi Y}\big] = \mathbb{E}\big[e^{i\xi(X_1+X_2+\cdots)}\big] = \mathbb{E}\big[e^{i\xi X_1} e^{i\xi X_2} \cdots\big] = \prod_{i\in\mathbb{N}} \mathbb{E}\big[e^{i\xi X_i}\big] = \prod_{i\in\mathbb{N}} \theta_{X_i}(\xi).$$

□

In this section, we introduced the measure-theoretic background required to represent random experiments through probability spaces and investigate particular outcomes using random variables. This will allow us to define stochastic processes, which are used in our models.

2.2 Stochastic processes

Stochastic processes are mathematical objects that are used to model phenomena that evolve randomly, such as stock prices. As we have seen, random variables may represent the outcome of a random experiment. Often, we want to investigate how this outcome varies over time and this is achieved by introducing stochastic processes.

Definition 17 (Stochastic process)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and (S, Σ) be a measurable space. Let T be some set, called the* ***index set****, and define an S-valued random variable X_t for each $t \in T$. A* ***stochastic process*** *is the collection of these random variables, $\{X_t \colon t \in T\}$. Note that all of the random variables are defined on the same measurable space.*

Given a filtration, $(\mathcal{F}_t)$, such that X_t is $\mathcal{F}_t$-measurable for all $t \in \mathbb{R}_+$, we say that the stochastic process is defined on the ***filtered probability space****, $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$.*

Historically, the index set, T, was a subset of $\mathbb{R}$, such as the natural numbers, $\mathbb{N}$, or an interval $[0, n]$, which allowed elements of the index set, $t \in T$, to be interpreted as time. While this is still often the case, the index set can be any totally ordered set. Another common choice for the index set is n-dimensional Euclidean space, where $t \in T$ can then be interpreted as a point in space. Sets that are not totally ordered may also be used, but many properties and theorems do not apply in this case.

A stochastic process, $\{X_t \colon t \in T\}$, is actually a function of two variables, $\omega \in \Omega$ and $t \in T$. Accordingly, it may also be written as $\{X(t, \omega) \colon t \in T, \omega \in \Omega\}$. Often, we want to fix $\omega \in \Omega$ and investigate a single outcome of the stochastic process, giving the function

$$X(\,\cdot\,, \omega) \colon T \to S,$$

which is called a *sample function* or *realisation* of the stochastic process. When T is interpreted as time, it is also called a **sample path** of the process.

Finally, if $S = \mathbb{R}$ or $S = \mathbb{Z}$, an **increment** of the stochastic process between the points $t_1, t_2 \in T$ is the difference between the underlying random variables at those points, $X_{t_2} - X_{t_1}$, which is itself another S-valued random variable. An increment represents how much the stochastic process changes between the two given points. This is usually considered when T is interpreted as time, where the increment is then the change in the process over a time period from t_1 to t_2.

Given a stochastic process, $\{X_t \colon t \in \mathbb{R}_+\}$, defined on probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the σ-field $\mathcal{F}$ corresponds to the amount of information that is contained in the stochastic process because it identifies which subsets of Ω have assigned probabilities. When we have a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, the filtration determines the amount of information contained in the stochastic process at time $t \in \mathbb{R}_+$. As more time passes, additional information is available, meaning that the filtration should become finer (have additional elements). This is why the σ-fields that form a filtration are ordered subsets.

If we do not have a filtered space, we can generate a filtration from the stochastic process, called the *natural filtration*.

Definition 18. *Let $\{X_t : t \in \mathbb{R}_+\}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For each $t \in \mathbb{R}^+$, let*

$$\mathcal{F}_t = \sigma(X_s \mid 0 \leq s \leq t)$$

*be the σ-field generated by all of the random variables X_s for $0 \leq s \leq t$. Then, $(\mathcal{F}_t)_{t\in\mathbb{R}^+}$ is the **natural filtration** of the stochastic process.*

Stochastic processes are mathematical objects that evolve according to a given probability distribution with random realisations and can be grouped into various categories, including random walks, martingales, Markov processes, Lévy processes, Gaussian processes, renewal processes, and branching processes. Two of the most famous and important examples of stochastic processes are Brownian motion and the Poisson process. Brownian motion was used by Louis Bachelier and others later to study financial markets, while the Poisson process was used by Agner Erlang to study the number of phone calls that occur in a certain period of time. Stochastic processes in general have found significant application in modelling systems and phenomena that appear to vary randomly, such as bacterial growth, electrical currents, and movement of gas molecules. In our research, we use the subcategory of stochastic processes called Lévy processes.

2.3 Lévy processes

Lévy processes are an important class of stochastic processes, named in recognition of French mathematician Paul Lévy, who played an instrumental role in the field. Throughout history, Lévy processes have been known by several different names. In fact, in 1940, Paul Lévy simply referred to them as a subclass of *processus additifs* (*additive processes*), which are processes that have independent increments. Later, in the 1960s and 70s, literature mostly referred to Lévy processes with the descriptive title of *processes with stationary and independent increments*, before the name Lévy process was eventually adopted. This subclass of stochastic processes is defined as follows.

Definition 19 (Lévy process)**.** *A stochastic process, $X = \{X_t : t \in \mathbb{R}_+\}$, on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a **Lévy process** if it satisfies the following.*

i. $X_0 = 0$ $\mathbb{P}$*-almost surely.*

ii. X has stationary increments:

$$X_t - X_s \overset{d}{=} X_{t-s} \quad \forall\, 0 \leq s \leq t.$$

iii. X has independent increments:

$$X_t - X_s \text{ is independent of } \{X_u \colon 0 \leq u \leq s\} \quad \forall\, 0 \leq s \leq t.$$

iv. Paths of X is càdlàg $\mathbb{P}$-a.s..

Lévy processes are an extremely rich class of processes with a wide range of properties. To better understand these processes, [17] introduced the notion of *infinitely divisible* distributions, and showed a one-to-one correspondence with Lévy processes, which we will see in Proposition 2.

Definition 20 (Infinitely divisible distribution)**.** *An $\mathbb{R}$-valued random variable, Θ, has an **infinitely divisible distribution** if, for each $n \in \mathbb{N}$, there exists a sequence of i.i.d. random variables, $\Theta_1, \ldots, \Theta_n$, such that*

$$\Theta \stackrel{d}{=} \Theta_1 + \cdots + \Theta_n.$$

Equivalently, the law, μ, of an $\mathbb{R}$-valued random variable is infinitely divisible if, for each $n \in \mathbb{N}$, there exists another random variable with law μ_n such that $\mu = \mu_n^{*n}$, where $\mu_n^{*n} = \mu_n * \mu_n * \cdots * \mu_n$ is the n-fold convolution of μ_n.

Proposition 2. *Let $X = \{X_t \colon t \in \mathbb{R}_+\}$ be a stochastic process. Then X is a Lévy process if and only if X_t has an infinitely divisible distribution for all $t \in \mathbb{R}_+$.*

The proof of this proposition requires several important objects that will be defined throughout this section. First, the forward direction of Proposition 2 can be shown directly from the previous definitions. Let $X = \{X_t \colon t \in \mathbb{R}_+\}$ be any Lévy process. For any $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$, write X_t as the following telescoping sum:

$$X_t = (X_{t/n} - X_0) + (X_{2t/n} - X_{t/n}) + \cdots + (X_t - X_{(n-1)t/n}).$$

The terms in this sum are non-overlapping intervals of equal lengths, so, by conditions (ii) and (iii) in Definition 19, they are each equal to the random variable $X_{t/n}$, independently. Hence, we have found a sequence of n i.i.d. random variables (n independent copies of $X_{t/n}$) such that

$$X_t = X_{t/n}^{(1)} + X_{t/n}^{(2)} + \cdots + X_{t/n}^{(n)}, \tag{2.1}$$

satisfying Definition 20 and showing that X_t has an infinitely divisible distribution.

To determine the reverse direction of Proposition 2, we introduce the *characteristic exponents* of random variables and Lévy processes. A random variable can be uniquely identified using its characteristic exponent, which is a function that completely describes the distributional properties of the variable.

Definition 21 (Characteristic exponent of random variable)**.** *The* ***characteristic exponent*** *of an $\mathbb{R}$-valued random variable, Θ, is a function $\Psi\colon \mathbb{R} \to \mathbb{C}$, given by*

$$\Psi(\xi) \coloneqq -\ln \mathbb{E}\big[e^{i\xi\Theta}\big].$$

Infinitely divisible random variables can be identified by their characteristic exponents, using the following lemma.

Lemma 3. *Let Θ be an $\mathbb{R}$-valued random variable with characteristic exponent Ψ. Then, Θ has an infinitely divisible distribution if and only if, for all $n \in \mathbb{N}$, there exists another characteristic exponent, Ψ_n, such that $\Psi(\xi) = n\Psi_n(\xi)$ for all $\xi \in \mathbb{R}$.*

Proof. See [38, Section 1.1]. □

We also have the characteristic exponent for a Lévy process, building the definition as follows. Suppose that we define the characteristic exponent of the Lévy process at time t in the same way as for a random variable (Definition 21), $\Psi_t(\xi) = -\ln \mathbb{E}\big[e^{i\xi X_t}\big]$. Then, using (2.1), $\Psi_t(\xi) = n\Psi_{t/n}(\xi)$. Applying (2.1) again, for another $m \in \mathbb{N}$, we have:

$$m\Psi_1(\xi) = \Psi_m(\xi) = n\Psi_{m/n}(\xi),$$

which implies that

$$\Psi_{m/n}(\xi) = \frac{m}{n}\Psi_1(\xi).$$

Let $t = m/n \in \mathbb{Q}$. So, for any rational $t > 0$,

$$\Psi_t(\xi) = t\Psi_1(\xi). \tag{2.2}$$

Additionally, recall that X is càdlàg, and thus, almost surely right continuous. By the dominated convergence theorem, this implies that $e^{-\Psi_t(\xi)}$ and $\Psi_t(\xi)$ are also right continuous for all $t \in \mathbb{R}_+$. Hence, for any irrational number, $t \in \mathbb{R}\backslash\mathbb{Q}$, $\Psi_t(\xi)$ can be calculated by choosing a decreasing sequence of rational numbers $\{t_n\colon n \in \mathbb{N}\}$ that converges to t (from above) as n tends to infinity. So, (2.2) holds for all $t \in \mathbb{R}_+$. Accordingly, we define the characteristic exponent of a Lévy process, $X = \{X_t\colon t \in \mathbb{R}_+\}$, to be the characteristic exponent of X_1, which must have an infinitely divisible distribution.

Definition 22 (Characteristic exponent of Lévy process)**.** *Let $X = \{X_t\colon t \in \mathbb{R}_+\}$ be a Lévy process. The* ***characteristic exponent*** *of X is a function $\Psi\colon \mathbb{R} \to \mathbb{C}$ such that*

$$\mathbb{E}\big[e^{i\xi X_t}\big] = e^{-t\Psi(\xi)},$$

or, equivalently,

$$\Psi(\xi) \coloneqq -\frac{\ln \mathbb{E}\big[e^{i\xi X_t}\big]}{t} = -\ln \mathbb{E}\big[e^{i\xi X_1}\big].$$

As with random variables, every Lévy process can be uniquely and completely specified by its characteristic exponent. Accordingly, studying the structure and properties of the characteristic exponent is extremely useful, as we will see later.

Finally, using the *Lévy-Khintchine formula*, we can show that the characteristic exponents of Lévy processes and infinitely divisible distributions are equivalent, and hence prove the reverse direction of Proposition 2. This theorem was derived in Lévy's 1937 paper, *Théorie de l'addition des variables aléatoires* [39] and introduces an expression for the characteristic exponent of an infinitely divisible distribution. We will see that this expression is identical to the characteristic exponent of a Lévy process.

Theorem 1 (Lévy-Khintchine formula for infinitely divisible distributions)**.** *A random variable with characteristic exponent Ψ has an infinitely divisible distribution if and only if there exists a triple (a, s, Π), where $a \in \mathbb{R}$, $s \in \mathbb{R}_+$, and Π is a measure concentrated on $\mathbb{R}\backslash\{0\}$ satisfying $\int_{\mathbb{R}}(1 \wedge x^2)\,\Pi(dx) < \infty$, such that*

$$\Psi(\xi) = -ia\xi + \frac{1}{2}s^2\xi^2 + \int_{\mathbb{R}} \left(1 - e^{i\xi x} + i\xi x \mathbb{1}_{\{|x|<1\}}\right) \Pi(dx).$$

Proof. See [40, Proposition 6.10]. □

The characteristic exponent of a random variable with an infinitely divisible distribution exactly defines a unique Lévy process with the same characteristic exponent. This is demonstrated by the Lévy-Khintchine formula for Lévy processes, which states that a Lévy process can be constructed with the same characteristic exponent as in Theorem 1. Note that, while the expressions in Theorems 1 and 2 are identical, the statements are not the same. Theorem 1 states that, for any infinitely divisible distribution, there is a triple (a, s, Π) such that the characteristic exponent has the required form. Alternatively, Theorem 2 states that, given any triple (a, s, Π), we can define a Lévy process with the required characteristic exponent. This proves that every infinitely divisible distribution corresponds to exactly one Lévy process, the reverse direction of Proposition 2.

Theorem 2 (Lévy-Khintchine formula for Lévy processes)**.** *Let $a \in \mathbb{R}$, $s \in \mathbb{R}_+$, and Π be a measure concentrated on $\mathbb{R}\backslash\{0\}$ such that $\int_{\mathbb{R}}(1 \wedge x^2)\,\Pi(dx) < \infty$, defining the triple (a, s, Π), which we call the* ***Lévy triple****. For each $\xi \in \mathbb{R}$, define*

$$\Psi(\xi) = -ia\xi + \frac{1}{2}s^2\xi^2 + \int_{\mathbb{R}} \left(1 - e^{i\xi x} + i\xi x \mathbb{1}_{\{|x|<1\}}\right) \Pi(dx).$$

Then, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which a Lévy process is defined having the characteristic exponent Ψ. We call a the ***drift****, s the* ***diffusivity****, and Π the* ***Lévy measure*** *of the defined Lévy process.*

The proof of this theorem is rewarding since it helps to understand the structure of Lévy processes. However, we will instead investigate a stronger version of the theorem, the *Lévy-Itô decomposition* in Section 2.3.3, which demonstrates the path structure explicitly.

2.3.1 Variation and jump activity

There are many properties that can be investigated when discussing Lévy processes; two that are particularly important in our research are the *path variation* and *jump activity*. In this section, we define these terms and present some techniques that may be used to calculate them. We first define path variation for a function.

Definition 23. *Let $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a function. Given a bounded interval $[a,b]$ and a partition $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$, the **variation** of the function f over $[a,b]$ with partition $\mathcal{P}$ is defined by*

$$V_{\mathcal{P}}(f,[a,b]) = \sum_{j=1}^{n} |f(t_j) - f(t_{j-1})|.$$

*The function has **bounded variation over** $[\mathbf{a},\mathbf{b}]$ if*

$$V(f,[a,b]) \coloneqq \sup_{\mathcal{P}} V_{\mathcal{P}}(f,[a,b]) < \infty.$$

*The function has **bounded variation** if it has bounded variation over all intervals $[a,b]$.*

Bounded variation of a function is adapted to bounded variation of a Lévy process with the same definitions except in an almost-sure sense.

Definition 24 (Bounded variation of Lévy process)**.** *Let $X = \{X_t\colon t \in \mathbb{R}_+\}$ be a Lévy process. X has bounded variation if $\{X_t(\omega)\}_t$ has bounded variation for almost all $\omega \in \Omega$.*

Given a Lévy process, the following lemma is often used to determine whether the path variation is bounded. This statement is given in [51, Theorem 21.9] and [38, Lemma 2.12].

Lemma 4. *A Lévy process with Lévy triple (a,s,Π) has paths of bounded variation if and only if*

i. $s = 0$.

ii. $\displaystyle\int_{\mathbb{R}} (1 \wedge |x|)\,\Pi(dx) < \infty$.

Given the condition $\int_{\mathbb{R}}(1 \wedge x^2)\,\Pi(dx) < \infty$ in Theorem 2, (ii) is equivalent to the statement, $\int_{|x|\le 1} |x|\,\Pi(dx) < \infty$.

Jump activity has a similar interpretation to path variation, but is specifically concerned with the jumps in the Lévy process, which are determined by the Lévy measure.

Definition 25. *Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a Lévy process. Then, X has finite **jump activity** if almost all paths have a finite number of jumps along any finite interval. Otherwise, the process has infinite jump activity.*

As with path variation, we have a result that provides a method to determine the jump activity of a given process, given in [16, Proposition 2].

Lemma 5. *Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a Lévy process with Lévy measure Π. Then, X has finite jump activity if and only if*

$$\Pi(\mathbb{R}) = \int_{\mathbb{R}} \Pi(dx) < \infty.$$

The condition for finite jump activity in Lemma 5 is stronger than the condition (ii) for bounded variation in Lemma 4. This means that pure jump processes (where $s = 0$ so that (i) in Lemma 4 is satisfied) with finite jump activity form a subclass of those with bounded variation. In other words, every jump process with finite jump activity must also have bounded variation. We verify this statement thoroughly next.

Lemma 6. *Let X be a pure jump Lévy process (diffusivity $s = 0$) with Lévy measure Π. If X has finite jump activity, X has bounded variation.*

Proof. Suppose that X, a pure jump process, has finite jump activity. By Lemma 5,

$$\Pi(\mathbb{R}) = \int_{\mathbb{R}} \Pi(dx) < \infty. \tag{2.3}$$

Additionally, recall from the Levy-Khintchine formula in Theorem 2, the Lévy measure satisfies

$$\int_{\mathbb{R}} (1 \wedge x^2)\, \Pi(dx) < \infty. \tag{2.4}$$

Since the integrands are both positive, (2.3) and (2.4) respectively imply that

$$\int_{|x|\geq 1} \Pi(dx) < \infty \quad \text{and} \quad \int_{|x|<1} |x|\, \Pi(dx) < \infty.$$

This implies

$$\int_{|x|<1} |x|\, \Pi(dx) + \int_{|x|\geq 1} \Pi(dx) = \int_{\mathbb{R}} (1 \wedge |x|)\, \Pi(dx) < \infty,$$

which is condition (ii) of Lemma 4. Hence, by Lemma 4, X has bounded variation. □

2.3.2 Examples

Now that we have defined Lévy processes and introduced some of their key properties, we will present several important examples of Lévy processes: Brownian motion, Poisson processes, compound Poisson processes, and subordinators.

Brownian motion

Perhaps the most notable Lévy process is the *Wiener process*, named after Norbert Wiener, who investigated many of its properties. However, Wiener processes are often called *Brownian motion* because they were historically used to model the random motion of particles suspended in a fluid, which is also known as Brownian motion. This process appears regularly in various fields such as pure and applied mathematics, quantitative finance, biology, and physics. Apart from being an effective model of many phenomena, Brownian motion can be used to generate and describe other stochastic processes. Stochastic models that are driven by Brownian motion are generally tractable and relatively easy to implement.

Brownian motion, $W = \{W_t : t \in \mathbb{R}_+\}$, is a stochastic process with almost-surely continuous paths, and increments that are independent, stationary, and normally distributed with mean μt and variance $\sigma^2 t$, for some $\mu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}_+$. When $\mu = 0$ and $\sigma = 1$, the process is called a *standard Brownian motion*. A simulation of two Brownian motions, with means $\mu = 0$ and $\mu = 1.4$ and both with variance $\sigma^2 = 4$, is shown in Figure 2.1. Below, we calculate the characteristic exponent of W and demonstrate that Brownian motion is a Lévy process, before discussing the path variation and jump activity.

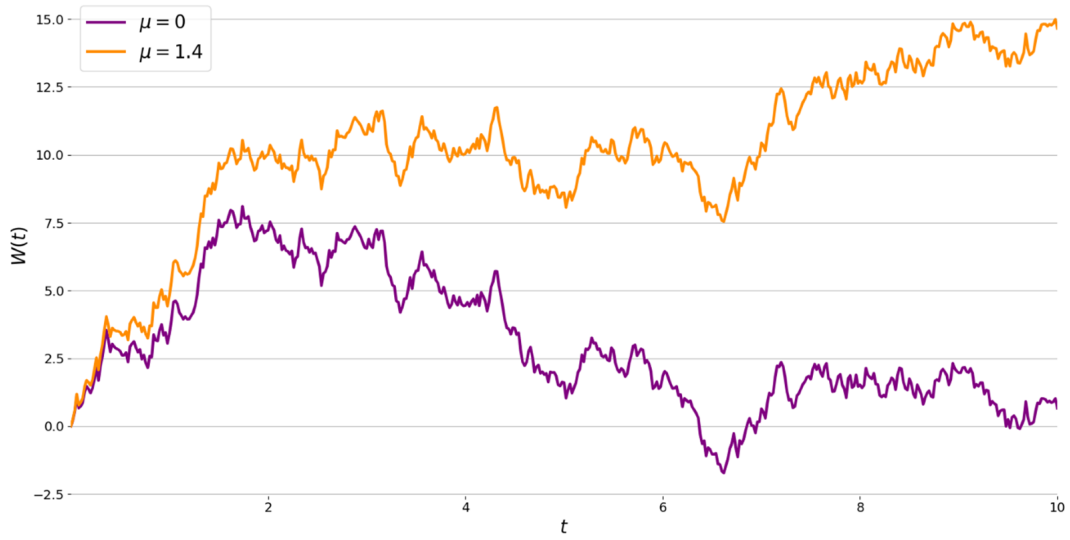


Figure 2.1: Brownian motion realisations with $\mu = 0$ and $\mu = 1.4$, both with $\sigma^2 = 4$.

Recall that a normally distributed random variable with mean a and variance s^2 has probability density function (law) $f(x) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2s^2}(x-a)^2}$. So, the characteristic function of a Brownian motion at time t, W_t, is

$$\begin{aligned}
\theta_{W_t}(\xi) &= \mathbb{E}\left[e^{i\xi W_t}\right] \\
&= \int_{\mathbb{R}} e^{i\xi x} \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{1}{2\sigma^2 t}(x-\mu t)^2}\, dx \\
&= \int_{\mathbb{R}} \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{1}{2\sigma^2 t}\left(x-(\mu-i\sigma^2\xi)t\right)^2} e^{-\frac{1}{2\sigma^2 t}\left(\mu^2t^2-(\mu t-i\sigma^2 t\xi)^2\right)}\, dx \\
&= e^{i\xi\mu t-\frac{\sigma^2 t\xi^2}{2}} \int_{\mathbb{R}} \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{1}{2\sigma^2 t}\left(x-(\mu-i\sigma^2\xi)t\right)^2}\, dx \\
&= e^{i\xi\mu t-\frac{\sigma^2 t\xi^2}{2}},
\end{aligned}$$

since, in the second to last expression, we have the integral of the probability density function of a $N(\mu t - i\sigma^2 t\xi,\, \sigma^2 t)$ variable over its support. Thus, by Definition 22, the characteristic exponent of a Brownian motion is

$$\Psi_W(\xi) = -\ln\left(\mathbb{E}\left[e^{i\xi W_1}\right]\right) = -\ln\left(e^{i\xi\mu-\frac{\sigma^2\xi^2}{2}}\right) = -i\xi\mu + \frac{\sigma^2\xi^2}{2}.$$

Observe that this expression satisfies the Lévy-Khintchine formula (Theorem 2) with $a = \mu$, $s = \sigma$, and $\Pi = 0$. This alone shows that W is a Lévy process, but we could also use Proposition 2 and show that W_t has an infinitely divisible distribution for every $t \in \mathbb{R}_+$. Let $n \in \mathbb{N}$ and define $Y_j^{(n)} \sim N\left(\frac{\mu t}{n}, \frac{\sigma^2 t}{n}\right)$ for $j = 1, \ldots, n$, independently. Let θ_Y be the characteristic function of Y_j, which is the same for all j since the variables have the same distribution. Similar to W,

$$\theta_Y(\xi) = e^{\frac{i\xi\mu t}{n}-\frac{\sigma^2 t\xi^2}{2n}}.$$

By Lemma 2,

$$\theta_{Y_1^{(n)}+Y_2^{(n)}+\cdots+Y_n^{(n)}}(\xi) = \prod_{j=1}^{n} \theta_Y = \prod_{j=1}^{n} e^{\frac{i\xi\mu t}{n}-\frac{\sigma^2 t\xi^2}{2n}} = e^{i\xi\mu t-\frac{\sigma^2 t\xi^2}{2}} = \theta_W(\xi).$$

Thus, by Lemma 1, $W_t \overset{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, as required for W_t to be infinitely divisible and hence, the Brownian motion is a Lévy process. Finally, it is also straightforward to show that Brownian motion satisfies the definition of a Lévy process (Definition 19). Properties (i), (ii), and (iii) are immediately satisfied, with the additional condition that the distribution of the increments must be normal. Since continuity is a stronger condition than càdlàg (every continuous process is also càdlàg), (iv) is also satisfied.

It is well-known (see, *e.g.*, [26, Exercise 9.8]) that Brownian motion has unbounded variation. Further, since the paths of Brownian motion are almost-surely continuous, it does not admit jumps and has finite (zero) jump activity.

Poisson process

The Poisson distribution is a discrete probability distribution that corresponds to the number of independent events in a fixed time interval (or distance, area, volume), for a given constant rate of occurrence. A Poisson process, $N = \{N_t : t \in \mathbb{R}_+\}$, is a stochastic process with almost-surely càdlàg paths and increments that are independent, stationary, and have a Poisson distribution with rate λt, for some $\lambda > 0$. Since the Poisson distribution is discrete, Poisson processes are strict jump processes, as demonstrated by the simulation in Figure 2.2. These processes are often used in queuing theory to model the arrival of customers, phone calls, or earthquakes. We will show that Poisson processes are Lévy processes, calculate their characteristic exponent, and discuss their variation and jump activity.

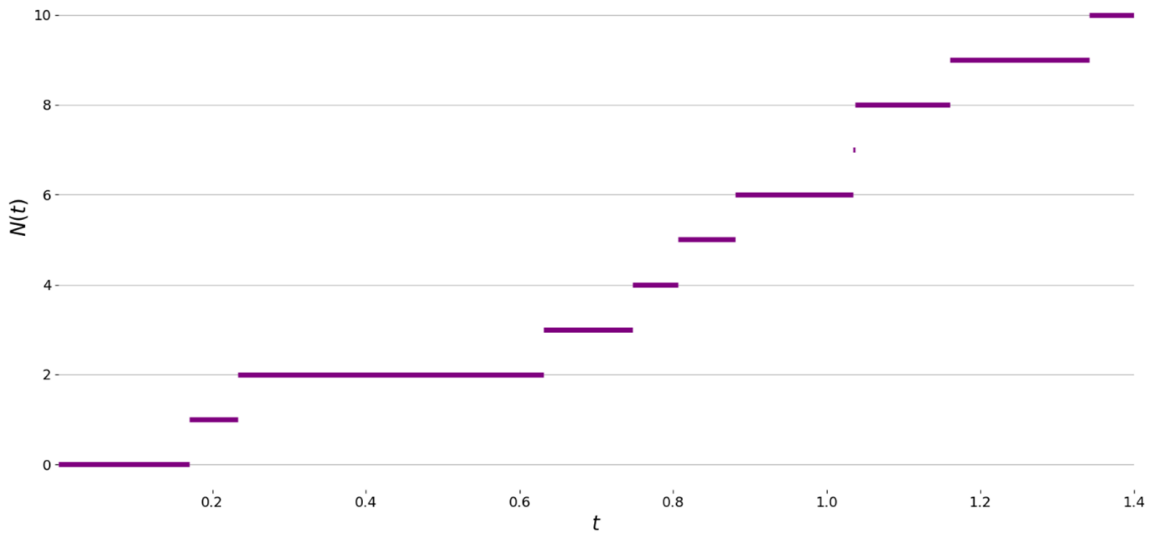


Figure 2.2: Poisson process realisation with rate $\lambda = 5$.

A Poisson random variable with rate γ has probability mass function $p(k) = \frac{\gamma^k e^{-\gamma}}{k!}$. So, using $e^x = \sum_{k\in\mathbb{N}} \frac{x^k}{k!}$, the characteristic function of a Poisson process at time t, N_t, is

$$\theta_{N_t}(\xi) = \mathbb{E}\left[e^{i\xi N_t}\right] = \sum_{k\in\mathbb{N}} e^{i\xi k} \frac{(\lambda t)^k e^{-\lambda t}}{k!} = e^{\lambda t(e^{i\xi}-1)}.$$

Thus, the characteristic exponent of a Poisson process is

$$\Psi_N(\xi) = -\ln\left(\mathbb{E}\left[e^{i\xi N_1}\right]\right) = -\ln\left(e^{\lambda(e^{i\xi}-1)}\right) = \lambda(1 - e^{i\xi}).$$

Observe that this expression satisfies the Lévy-Khintchine formula with $a = s = 0$, and $\Pi = \lambda\delta_1$, where δ_1 is the Dirac measure supported on $\{1\}$. As before, we show that N_t has an infinitely divisible distribution. Let $n \in \mathbb{N}$ and define $Y_j^{(n)} \sim \text{Poisson}\left(\frac{\lambda t}{n}\right)$. Let θ_Y be the characteristic function of Y_j, which is the same for all j. Similar to N,

$$\theta_Y(\xi) = e^{\frac{\lambda t}{n}(e^{i\xi}-1)}.$$

By Lemma 2,

$$\theta_{Y_1^{(n)}+Y_2^{(n)}+\cdots+Y_n^{(n)}}(\xi) = \prod_{j=1}^{n} \theta_Y = \prod_{j=1}^{n} e^{\frac{\lambda t}{n}(e^{i\xi}-1)} = e^{\lambda t(e^{i\xi}-1)} = \theta_N(\xi).$$

Thus, by Lemma 1, $N_t \overset{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, as required for N_t to be infinitely divisible and hence, Poisson processes are Lévy processes. Again, the Poisson process also satisfies Definition 19 with the additional requirement that the increments have Poisson distributions.

Since the support of the Poisson distribution is $\mathbb{Z}_+$, the change in a Poisson process over any increment is non-negative. Thus, Poisson processes are non-decreasing, meaning that they have bounded variation over finite time horizons [38]. Since the Poisson process is a strict jump process, by Lemma 6, this implies that the jump activity is bounded also.

We have already seen two Lévy processes that are vastly different in the continuity of their paths, distribution of their increments, and their boundedness. This begins to show the richness of the class of Lévy processes.

Compound Poisson process

A compound Poisson process, $D = \{D_t : t \in \mathbb{R}_+\}$, is a stochastic process that allows jumps in its paths whose frequency is according to a Poisson process, $\{N_t : t \in \mathbb{R}_+\}$, and magnitude is according to some other distribution, F. In other words,

$$D_t = \sum_{j=1}^{N_t} A_j,$$

where $N_t \sim \text{Poisson}(\lambda t)$ and $A_j \sim F$, independently for each j. Compound Poisson processes have a similar path structure to Poisson processes shown in Figure 2.2, but with different jump sizes. Again, we show that compound Poisson processes are Lévy processes and calculate their characteristic exponent.

By the law of total probability, the characteristic function of a compound Poisson process at time t, D_t, is

$$\begin{aligned}\theta_{D_t}(\xi) &= \mathbb{E}\Big[e^{i\xi \sum_{j=1}^{N_t} A_j}\Big] \\ &= \sum_{n\in\mathbb{N}} \mathbb{E}\Big[e^{i\xi \sum_{j=1}^{N_t} A_j} \;\Big|\; N_t = k\Big]\, \mathbb{P}(N_t = k) \\ &= \sum_{n\in\mathbb{N}} \mathbb{E}\Big[e^{i\xi \sum_{j=1}^{k} A_j}\Big]\, \frac{(\lambda t)^k e^{-\lambda t}}{k!}.\end{aligned}$$

As A_j are i.i.d.,

$$\begin{aligned}\theta_{D_t}(\xi) &= \sum_{n\in\mathbb{N}} \mathbb{E}\left[e^{i\xi A_1}\right]^k \frac{(\lambda t)^k e^{-\lambda t}}{k!} \\ &= \sum_{n\in\mathbb{N}} \left(\int_{\mathbb{R}} e^{i\xi x} F(dx)\right)^k \frac{(\lambda t)^k e^{-\lambda t}}{k!} \\ &= e^{-\lambda t\left(\int_{\mathbb{R}} (1-e^{i\xi x}) F(dx)\right)}.\end{aligned}$$

Thus, the characteristic exponential of a compound Poisson process is

$$\Psi_D(\xi) = -\log \mathbb{E}\Big[e^{i\xi \sum_{j=1}^{N_1} A_j}\Big] = \lambda \left(\int_{\mathbb{R}} (1 - e^{i\xi x}) F(dx)\right).$$

Observe that this satisfies the Lévy-Khintchine formula with $a = \lambda \int_{\mathbb{R}} x \mathbb{1}_{\{|x|<1\}} F(dx)$, $s = 0$, and $\Pi(dx) = \lambda F(dx)$. Further, with similar reasoning to that seen for Brownian motion and Poisson processes, D_t can be shown to have an infinitely divisible distribution. Thus, compound Poisson processes are Lévy processes.

Subordinators

A subordinator is a Lévy process that is non-negative, which, due to the stationary increments property of Lévy processes, implies that they must be non-decreasing almost-surely. Clearly, subordinators must have non-negative drift, no diffusion component, and only positive jumps of finite variation. Thus, the characteristic exponent of a subordinator is given by

$$\Psi(\xi) = -i\mu\xi + \int_{\mathbb{R}_+} \left(1 - e^{i\xi x}\right) \Pi(dx). \tag{2.5}$$

Given a stochastic process, subordinators are often used in the theory of local time to generate a new stochastic process, called the *subordinated process*, where the evolution of

time is according to the subordinator [25]. The number of time steps that occur for the subordinated process in a given unit of chronological time is determined by the value of the subordinator in that increment.

2.3.3 Lévy-Itô decomposition

The Lévy-Itô decomposition provides insight into the path structure of a general Lévy process by breaking it into three other Lévy processes, each with a particular path type. We summarise the theorem to provide additional understanding of the paths of the Lévy processes that we have seen. A more detailed statement of the theorem is discussed in [38]. Recall the Lévy-Khintchine formula in Theorem 2, which states that the characteristic exponent of a Lévy process has the form:

$$\Psi(\xi) = -ia\xi + \frac{1}{2}s^2\xi^2 + \int_{\mathbb{R}} \left(1 - e^{i\xi x} + i\xi x \mathbb{1}_{\{|x|<1\}}\right) \Pi(dx).$$

Consider splitting this formula into three components,

$$\begin{aligned}\Psi(\xi) = &\left\{-ia\xi + \frac{1}{2}s^2\right\} \\ &+ \left\{\Pi(\mathbb{R}\backslash(-1,1)) \int_{|x|\geq 1} \left(1 - e^{i\xi x}\right) \frac{\Pi(dx)}{\Pi(\mathbb{R}\backslash(-1,1))}\right\} \\ &+ \left\{\int_{|x|<1} \left(1 - e^{i\xi x} + i\xi x\right) \Pi(dx)\right\}.\end{aligned}$$

Label each of the decomposed components,

$$\begin{aligned}\Psi^{(1)}(\xi) &= -ia\xi + \frac{1}{2}\sigma^2, \\ \Psi^{(2)}(\xi) &= \Pi(\mathbb{R}\backslash(-1,1)) \int_{|x|\geq 1} \left(1 - e^{i\xi x}\right) \frac{\Pi(dx)}{\Pi(\mathbb{R}\backslash(-1,1))}, \\ \Psi^{(3)}(\xi) &= \int_{|x|<1} \left(1 - e^{i\xi x} + i\xi x\right) \Pi(dx).\end{aligned}$$

Then, $\Psi^{(1)}$ is the characteristic exponent of a Brownian motion with drift a and diffusivity σ^2. If $\Pi(\mathbb{R}\backslash(-1,1)) = 0$, then $\Psi^{(2)} = 0$, which corresponds to a process that is identically 0. Otherwise, $\Psi^{(2)}$ is the characteristic exponent of a compound Poisson process where N_t has rate $\Pi(\mathbb{R}\backslash(-1,1))$ and A_j are i.i.d. with distribution $\Pi(dx)/\Pi(\mathbb{R}(-1,1))$ concentrated on $\{x\colon |x| \geq 1\}$. Additionally, according to [38, Theorem 2.10], $\Psi^{(3)}$ is also the characteristic exponent of a Lévy process, a *square-integrable martingale* [38]. These three processes are independent, such that the Lévy-Khintchine formula in Theorem 2 is the sum of three independent Lévy processes. It can be observed directly from the

definition of a Lévy process that this sum is also a Lévy process. This decomposition is called the Lévy-Itô decomposition, credited to Paul Lévy and Kiyosi Itô, and proves the Lévy-Khintchine formula in Theorem 2. It also shows that any Lévy process can be expressed as the sum of a Brownian motion, compound Poisson process, and a square-integrable martingale, which helps to understand the characteristics and path properties of a Lévy process.

2.3.4 Wiener-Hopf factorisation

Wiener-Hopf factorisation refers to a collection of results on the characteristics of the *ascending and descending ladder* processes of a stochastic process. It was initially developed by Norbert Wiener and Eberhard Hopf to solve integral equations, but has since been used in solving various types of differential equations. In general, the technique aims to define a pair of functions that are respectively analytic in the upper and lower complex half-planes and describe the equations and boundary conditions of the problem.

Varying expressions for the Wiener-Hopf factorisation formula exist within different disciplines. In modern probability, the formula relates the characteristic functions of a Lévy process to its supremum and infimum processes when observed at exponentially distributed times, or equivalently, at the arrival times of a Poisson process. Throughout this section and the entire thesis, $\mathbf{e}_q \sim \text{Exp}(q)$ will denote an exponentially distributed random variable that is independent of our Lévy processes. The supremum and infimum processes are defined as follows.

Definition 26 (Supremum and infimum processes)**.** *Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a stochastic process. The **supremum process** of X, $\overline{X} = \{\overline{X}_t : t \in \mathbb{R}_+\}$, is defined by*

$$\overline{X}_t \coloneqq \sup_{0 \le s \le t} X_s,$$

*and the **infimum process** of X, $\underline{X} = \{\underline{X}_t : t \in \mathbb{R}_+\}$, is defined by*

$$\underline{X}_t \coloneqq \inf_{0 \le s \le t} X_s.$$

First, given stochastic process X and $t \in \mathbb{R}_+$, the pairs

$$(\overline{X}_t, \overline{X}_t - X_t) \quad \text{and} \quad (X_t - \underline{X}_t, -\underline{X}_t) \tag{2.6}$$

have the same distribution under $\mathbb{P}$ [38, Lemma 3.5]. Additionally, if $X = \{X_t : t \in \mathbb{R}_+\}$ is a Lévy process, then it can be shown $\overline{X}_{\mathbf{e}_q}$ and $X_{\mathbf{e}_q} - \overline{X}_{\mathbf{e}_q}$ are independent and infinitely divisible [51]. This relationship allows us to prove the Wiener-Hopf factorisation formula.

Theorem 3 (Wiener-Hopf factorisation)**.** *Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a Lévy process. The **Wiener-Hopf factorisation** of X is*

$$\mathbb{E}\left[e^{i\xi X_{\mathbf{e}_q}}\right] = \mathbb{E}\left[e^{i\xi \overline{X}_{\mathbf{e}_q}}\right]\mathbb{E}\left[e^{i\xi \underline{X}_{\mathbf{e}_q}}\right], \quad \forall\, \xi \in \mathbb{R}.$$

Proof. Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a Lévy process and $\xi \in \mathbb{R}$. Then,

$$\mathbb{E}\Big[e^{i\xi X_{\mathbf{e}_q}}\Big] = \mathbb{E}\Big[e^{i\xi\left(X_{\mathbf{e}_q} - \overline{X}_{\mathbf{e}_q} + \overline{X}_{\mathbf{e}_q}\right)}\Big] = \mathbb{E}\Big[e^{i\xi\overline{X}_{\mathbf{e}_q}} e^{i\xi\left(X_{\mathbf{e}_q} - \overline{X}_{\mathbf{e}_q}\right)}\Big].$$

Since $\overline{X}_{\mathbf{e}_q}$ and $X_{\mathbf{e}_q} - \overline{X}_{\mathbf{e}_q}$ are independent,

$$\mathbb{E}\Big[e^{i\xi X_{\mathbf{e}_q}}\Big] = \mathbb{E}\Big[e^{i\xi\overline{X}_{\mathbf{e}_q}}\Big]\mathbb{E}\Big[e^{i\xi\left(X_{\mathbf{e}_q} - \overline{X}_{\mathbf{e}_q}\right)}\Big].$$

Finally, using (2.6),

$$\mathbb{E}\Big[e^{i\xi X_{\mathbf{e}_q}}\Big] = \mathbb{E}\Big[e^{i\xi\overline{X}_{\mathbf{e}_q}}\Big]\mathbb{E}\Big[e^{i\xi\underline{X}_{\mathbf{e}_q}}\Big].$$

□

The *Wiener-Hopf factors* are defined to be the two terms on the right-hand side of the Wiener-Hopf factorisation.

Definition 27 (Wiener-Hopf factors)**.** *Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a Lévy process. The* ***Wiener-Hopf factors*** *of X are:*

$$\phi_q^+(\xi) := \mathbb{E}\Big[e^{i\xi\overline{X}_{\mathbf{e}_q}}\Big], \quad \phi_q^-(\xi) := \mathbb{E}\Big[e^{i\xi\underline{X}_{\mathbf{e}_q}}\Big].$$

Currently, there are relatively few Lévy processes for which we can calculate explicit expressions for Wiener-Hopf factors and this is an important problem in the literature. Standard Brownian motion is one example of a process for which we have known Wiener-Hopf factors, with the following form [38]:

$$\phi_q^+(\xi) = \frac{\sqrt{2q}}{\sqrt{2q} + i\xi}, \quad \phi_q^+(\xi) = \frac{\sqrt{2q}}{\sqrt{2q} - i\xi}.$$

The Wiener-Hopf factorisation formula involves the characteristic functions of the Lévy process and its supremum and infimum processes. Similar to the characteristic exponent, the characteristic function completely specifies a Lévy process, so expressions given by the Wiener-Hopf factorisation are extremely useful in investigating the dynamics of a Lévy process. Additionally, the characteristic function evaluated at an exponential random variable can be calculated analytically to derive an expression in terms of the characteristic exponent of the Lévy process. Using the tower property and then the definition of expectation,

$$\begin{aligned}
\mathbb{E}\Big[e^{i\xi X_{\mathbf{e}_q}}\Big] &= \mathbb{E}\Big[\mathbb{E}\big[e^{i\xi X_t} \;\big|\; \mathbf{e}_q = t\big]\Big] \\
&= \int_{\mathbb{R}_+} \mathbb{E}\big[e^{i\xi X_t}\big]\, qe^{-qt}\, dt \\
&= q\int_{\mathbb{R}_+} \mathbb{E}\big[e^{i\xi X_t}\big]\, e^{-qt}\, dt.
\end{aligned}$$

Now, recall, from Definition 22, that $\mathbb{E}\left[e^{i\xi X_t}\right] = e^{-\Psi(\xi)t}$, where $\Psi(\xi)$ is the characteristic exponent of X. Then,

$$\begin{aligned}\mathbb{E}\left[e^{i\xi X_{\mathbf{e}_q}}\right] &= q\int_{\mathbb{R}_+} e^{-\Psi(\xi)t}\, e^{-qt}\, dt \\ &= \frac{q}{q+\Psi(\xi)}\int_{\mathbb{R}_+} \big(q+\Psi(\xi)\big)\, e^{-(q+\Psi(\xi))t}\, dt \\ &= \frac{q}{q+\Psi(\xi)}. \end{aligned} \tag{2.7}$$

This new expression for the characteristic function is easily implemented since we often have simple formulas for the characteristic exponent, $\Psi(\xi)$, from the Lévy-Khintchine formula in Theorem 2. Using (2.7), the Wiener-Hopf factorisation formula in Theorem 3 can be written

$$\frac{q}{q+\Psi(\xi)} = \phi_q^+(\xi)\,\phi_q^-(\xi). \tag{2.8}$$

To conclude this section, we make a final observation that has several useful consequences: $\frac{q}{q+\Psi(\xi)}$, $\phi_q^+(\xi)$, and $\phi_q^-(\xi)$ are the characteristic functions of $X_{\mathbf{e}_q}$, $\overline{X}_{\mathbf{e}_q}$, and $\underline{X}_{\mathbf{e}_q}$, respectively. This follows immediately by Equation (2.8) and Definition 27, along with Definition 16. There are many well-known properties and theorems for characteristic functions which are available, such as the inversion formula.

Theorem 4. *Let Y be an $\mathbb{R}$-valued random variable with probability density function f_Y and characteristic function θ, which is integrable,*

$$\int_{\mathbb{R}} |\theta(y)|\, dy < \infty.$$

*Then, the **inversion formula** for Y states that*

$$f_Y(y) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iy\xi}\theta(\xi)\, d\xi. \tag{2.9}$$

Proof. See [46, Section 11.2]. □

Throughout this chapter, we have discussed the relevant areas of measure theory that are used to define stochastic processes and the related theorems that will later contribute to our research.

Chapter 3

Financial background

Mathematical finance is an area of applied mathematics that aims to understand and model financial markets. The first scholarly work in mathematical finance is considered to be Louis Bachelier's doctoral thesis, *Theory of Speculation*, in 1900 [2]. This is often marked as the origin of quantitative finance. However, the discipline officially emerged in the 1970s, following the work of Fischer Black, Myron Scholes, and Robert Merton in their influential papers, *The Pricing of Options and Corporate Liabilities* [5] and *Theory of Rational Option Pricing* [42]. These papers introduced the *Black-Scholes formula* and motivated further works in the field.

Our research is in mathematical finance, investigating a model that could be thought of as a descendent of the Black-Scholes formula. In this chapter, we set the scene for the thesis by delving into the financial background of the project. We discuss financial markets and the different objects that can be modelled, along with the properties that we want to achieve with these models. The concept of arbitrage is also introduced, and finally, the original Black-Scholes formula is discussed.

3.1 Financial markets and securities

Financial markets encompass a variety of institutions that generally aim to facilitate the flow of capital from those who have funds and those who need funds. This transfer occurs by the trading of *financial assets* (*financial instruments*), which refers to any resource that is expected to provide future benefits and can be categorised into tangible and intangible assets. Tangible assets, such as buildings, land, and machinery, have physical value, while intangible assets are legal contracts with potential future economic benefits, including patents, copyrights, and trademarks. Instruments such as stocks, bonds, and futures are also intangible assets, but are often referred to as securities [15]. A financial market is a destination where financial instruments may be exchanged [15, 53]. This may be a physical location, such as the New York Stock Exchange, or a virtual system such as

the NASDAQ. The first financial market, called the Bourse, opened in Belgium in 1531 and focused mostly on tangible assets [44]. In 1611, the Dutch East India Company was the first to offer stock trading [21]. Stock exchanges have grown significantly; most large companies today offer public stock trading. An example of a financial security, in the form of the TSLA stock over a 12-month period during 2021, is shown in Figure 3.1.

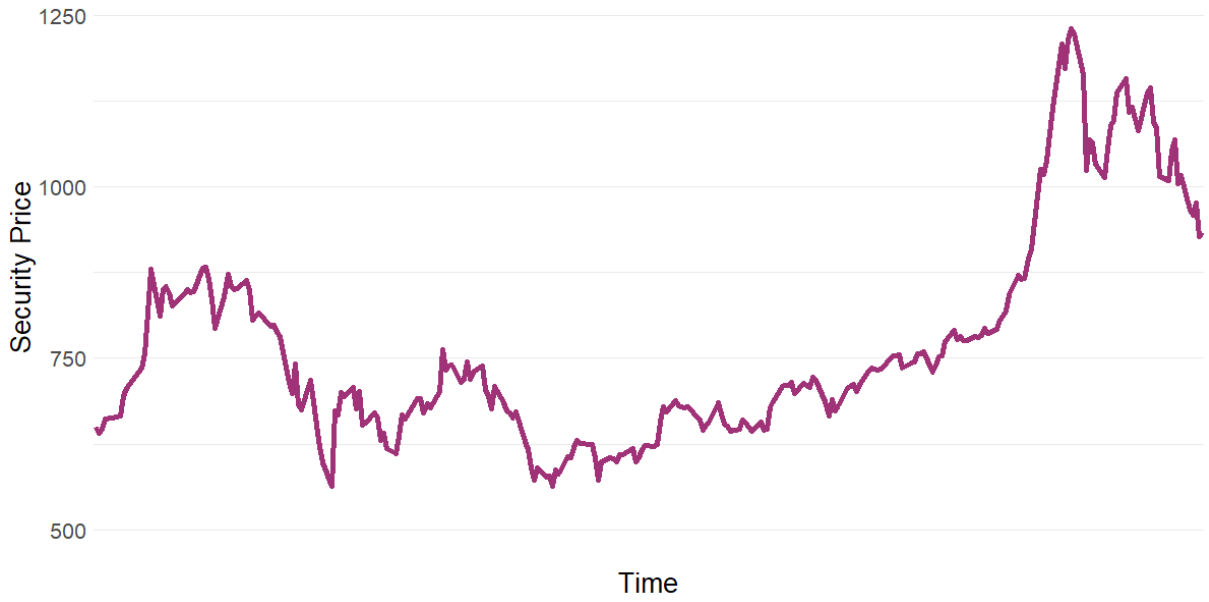


Figure 3.1: TSLA stock price evolution over the 12-month period in 2021.

The set of all financial objects that an investor owns is often called a *trading strategy*. These strategies should depend only on the current and previous market information, and are called *admissible* in this case. Additionally, a trading strategy is called *self-financing* if there is no external input or output of funds. These definitions are useful for discussing trading strategies and categorising financial markets [8, 48, 55].

Definition 28 (Trading strategy)**.** *A trading strategy,* $\chi = \{\chi_t \colon t \in \mathbb{R}_+\}$*, is a stochastic process, where* χ_t *represents the number of units of a financial security that an investor owns at time* t*.*

Definition 29 (Admissible trading strategy)**.** *A* ***trading strategy****,* $\chi = \{\chi_t \colon t \in \mathbb{R}_+\}$*, is called* ***admissible*** *if* χ *is a predictable process; that is,* χ *is measurable with respect to the* σ*-field on* $\mathbb{R}_+ \times \mathbb{R}_+$ *generated by all left-continuous adapted processes.*

Definition 30 (Self-financing strategy)**.** *An admissible trading strategy is called* ***self-financing*** *if no funds are withdrawn or added to the strategy, meaning the total value of*

the strategy, denoted $V_t(\chi)$, is constant over time. So, for any $s, t \in \mathbb{R}_+$,

$$V_t(\chi) \coloneqq \chi_s S_s = \chi_t S_t,$$

where S_t is the price of the security at time t, so that $\chi_t S_t$ is the number of units of the security held times the price of the security, the total value of the strategy.

When modelling financial markets, we want to use an *equivalent martingale measure* (EMM). An exact definition of this mathematical object will be introduced in Section 3.2.2. Heuristically, an EMM is a probability measure such that the expected value of the stock price under this measure is the fair price of the stock when adjusted for the involved risk. Of course, this is a desirable property, and thus, we want an EMM to exist.

3.1.1 Derivatives

A *derivative* is a financial object whose value depends on some underlying security, which may itself be a derivative. Some examples of these are *options*, *futures contracts*, and *swaps*; we will focus on options. Financial markets formed by derivatives are called *secondary markets*.

Option contracts are financial derivatives that provide the right, but not the obligation, to buy or sell an underlying security at a certain price and date, called the *strike price* and *maturity date*. The current price of the underlying security when the option is purchased is called the *spot price*. Options provide opportunities for advanced hedging and speculation, and limit potential losses. An option that allows the security to be purchased is a *call option*, while those where the security is sold are *put options*. Accordingly, the payouts for a call and put option are

$$G_{\text{call}}(t) = (S_t - K)^+ \quad \text{and} \quad G_{\text{put}}(t) = (K - S_t)^+, \tag{3.1}$$

respectively, where K is the strike price, S_t is the price of the underlying security at time t, and the function $(\,\cdot\,)^+ = \max(\,\cdot\,, 0)$. These basic contracts are sometimes called *vanilla options*. Examples of the structure of a put option are shown in Figure 3.2, with maturity date by T and two different strike prices, K. The underlying security for this option is the TSLA stock from Figure 3.1. Since we are considering a put option, the payout function is positive whenever the security price is less than the strike price. So, the owner could exercise the option in Figure 3.2a at any time and have a positive return. On the other hand, the option in Figure 3.2b has positive payout only at certain times. There are many different types of options; some options would allow the owner to exercise the contract during these periods of positive payout, while others may not. *American* and *European* options are the most common types that provide this distinction. An American option can be exercised at any time up to maturity, whereas a European option can only be exercised at maturity.

(a) Security price always lower than the strike. (b) Security price fluctuates around the strike.

Figure 3.2: Examples of potential put option contracts. Green and red indicate periods with positive and zero payout, respectively.

Option contracts offer several advantageous properties to both experienced traders and individual investors; they

- Provide increased cost-efficiency. Option fees are significantly less than purchasing the underlying security, meaning that less capital is required to invest in options.
- Limit the risk. Options can be less risky since the capital commitment is reduced.
- Allow for more variety in investment strategies. Options are extremely flexible and can be combined and hedged in many ways, allowing potential profit in any market condition.

Despite some initial skepticism, options have grown substantially in popularity due to these benefits.

Options originated in the form of markets called *bucket shops* in the 1920s. At this time, financial markets were largely mysterious and could not be utilised effectively. The introduction of the first explicit option pricing models, such as the Black-Scholes formula, allowed fair values to be assigned to financial objects, meaning that they could be effectively traded. This was the catalyst for the development of derivative markets around the world, which are a multi-trillion-dollar industry today. Derivative pricing formulae and algorithms have produced massive economic growth. Improved accuracy and speed in these models would allow the market to become fairer and more effective, showing the importance of continual development in this field. On the other hand, inaccuracies in these models, or their application in the wrong market conditions where the assumptions of the model are not satisfied, can have dire consequences. This further stresses the importance of producing accurate models and advocates for the development of new models, with different assumptions that may be applied more broadly.

3.1.2 Efficient and complete markets

There are several ways that financial markets are categorised based on their properties. Two important distinctions are *efficient* and *complete* markets [4, 8, 48, 55]. These classifications allow us to determine the existence and uniqueness of an equivalent martingale measure.

Market efficiency is determined by the presence of arbitrage, which is an opportunity for an investor to find undervalued or overvalued securities and beat the market by achieving a risk-free profit. The equations in Definition 31 state that an arbitrage opportunity is identified by a portfolio that has no initial value, but has guaranteed positive returns.

Definition 31 (Arbitrage opportunity)**.** *An arbitrage opportunity is a self-financing, admissible strategy χ with, $V_0(\chi) = 0$, $V_T(\chi) \geq 0$, and $\mathbb{E}[V_T(\chi)] > 0$.*

If all of the available financial information is reflected in the price of securities, there will be no arbitrage opportunities since the price will always be fair. This defines an efficient market [19].

Definition 32 (Efficient market)**.** *A financial market is called efficient if there are no arbitrage opportunities.*

In an efficient market, we can always find an equivalent martingale measure [7, 14]. This result is often called the first fundamental theorem of asset pricing [19], and may be proved using the Hahn-Banach separation theorem (see [4, Proposition 4.2.3]).

Theorem 5 (First fundamental theorem of asset pricing)**.** *A financial market is efficient if and only if there is at least one equivalent martingale measure.*

In the 1960s, Eugene Fama proposed the Efficient Market Hypothesis, which states that financial markets are approximately efficient [15, 53, 55]. Empirical studies have supported this claim, and accordingly, we should assume an efficient market when dealing with financial models.

Complete markets are those where, for every derivative in the market, there is a trading strategy that has the same value at maturity. A derivative is called *attainable* if we have this property [19]. The corresponding trading strategy is called a *replicating strategy* because it replicates the derivative.

Definition 33 (Attainable derivative)**.** *A derivative X with maturity T is called attainable if there is an admissible, self-financing trading strategy χ such that $V_T(\chi) = X$.*

Definition 34 (Complete market)**.** *A financial market is called complete if every derivative is attainable.*

This leads to the second fundamental theorem of asset pricing, which states that market completeness means that we have a unique equivalent martingale measure. The proof is given in [4, Section 4.3].

Theorem 6 (Second fundamental theorem of asset pricing)**.** *A market is both efficient and complete if and only if there is exactly one equivalent martingale measure.*

Many financial models, particularly earlier models, assume that the market is complete. However, empirical studies have shown markets are generally incomplete, so this is often an unreasonable assumption [52]. Probably the most famous option pricing formula, and an example of a model which does assume a complete market, is the Black-Scholes formula.

3.1.3 Barrier options

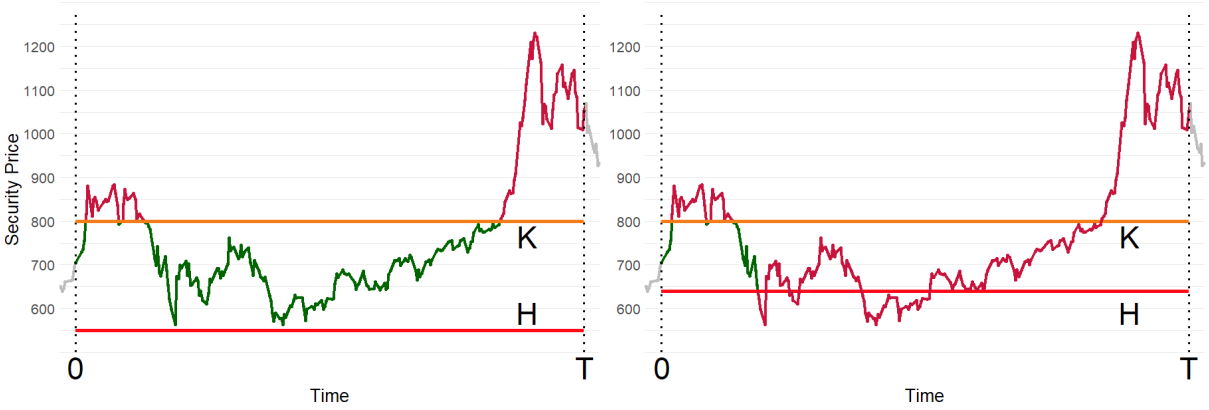
Further to the call and put distinction, option contracts vary significantly in how and when they can be exercised, and in the subsequent payout. As mentioned previously, the predominant subclasses of options are the American and European variations, which vary in the time that owners can exercise the contract. An American option may be exercised at any time up to maturity, while a European option can be exercised only at maturity. There are various other types of options, which may have different payment structures, expiration dates, and strike prices. These are often called *exotic options*; some examples are lookback, compound, and barrier options. The variety of option types leads to unique and interesting behaviours. Pricing different exotic options often requires completely different modelling techniques, so the literature in this area is vast.

Barrier options are a type of exotic option; they differ from vanilla options in the payment structure. Along with the standard parameters, a barrier option contract specifies the value of a barrier, H, and an additional categorisation as either *knock-in* or *knock-out*. A knock-in barrier option has no value until the barrier is reached, after which the payout function is the same as a vanilla option in (3.1). Adversely, a knock-out option has the same payout function as a vanilla option until the barrier is reached, at which point it becomes worthless. Accordingly, the payout functions for knock-in and knock-out barrier put options are

$$G_{\text{in}}(t) = (K - S_t)^+ \mathbb{1}\{\underline{S}_t < H\} \quad \text{and} \quad G_{\text{out}}(t) = (K - S_t)^+ \mathbb{1}\{\underline{S}_t > H\}. \tag{3.2}$$

In some cases, the owner of a knock-out option may be entitled to a partial refund, called a *rebate*, if it expires worthless. Two examples of potential knock-out barrier options on the TSLA stock from Figure 3.1, with strike K, maturity T, and barrier H, are shown in Figure 3.3. In Figure 3.3a, the security does not reach the barrier and so, the option payout is identical to that in Figure 3.2b. Contrarily, in Figure 3.3b the barrier is reached by the security, after which the option has zero payout, regardless of the strike price.

Aside from providing additional variety in investment strategies, barrier options also allow the seller to limit their risk and have lower premiums than vanilla options.



(a) Security does not reach the barrier.

(b) Security is deactivated by the barrier.

Figure 3.3: Examples of potential barrier put option contracts. Green and red indicate periods with positive and zero payout, respectively.

3.1.4 Bonds

Bonds are another type of financial security that should be considered when modelling the value of derivatives. Like other derivatives, bonds have a set maturity date, but they do not have a strike price. Instead, the issuer of a bond is obligated to repay the principal along with a specified amount of interest. Bonds are referred to as *fixed-income instruments* because there is a guaranteed rate of interest on the initial investment, although this rate can actually be variable. Companies, municipalities, states, and sovereign governments often issue bonds to finance projects and operations.

The guaranteed interest rate of bonds in a financial market provides a baseline return, which we call the *risk-free interest rate* of the market. If the expected return of a stock is less than the risk-free interest rate, then the owner is accepting unnecessary risk. Thus, the calculated return in financial models is scaled by the risk-free interest rate.

3.2 Stochastic modelling of options

Stochastic modelling is a form of mathematical modelling where some phenomenon is assumed to occur according to a stochastic process. Applying these models to price options is a prevalent problem in financial mathematics. Since the security price should be non-negative, we often use exponential stochastic processes, such that the security price is given by $S_t = S_0 e^{X_t}$, where $X = \{X_t : t \in \mathbb{R}_+\}$ is a stochastic process.

The literature presents a wide range of potential stochastic processes for the model, aiming to effectively satisfy several assumptions about financial markets and properties of the paths of security prices (see [7]). First, markets are often assumed to be efficient and incomplete, which we will discuss further in the next section. Empirical studies have shown that the log-distribution of security prices usually has fat tails. Therefore, our model should also exhibit fat tails in its path increments. Additionally, large sudden changes are relatively common in financial markets. Stochastic processes with continuous paths do not model this phenomenon effectively; thus, the literature has aimed to implement processes that admit jumps in their path (see, *e.g.*, [22, 28, 32]). Finally, studies suggest that there are options that are better modelled with processes of finite variation, and others where infinite variation is more appropriate [13, 15, 20, 53]. Hence, it is beneficial for models to allow for both of these possibilities.

Option price models must be fitted to the dynamics of the underlying security by considering two important properties that describe its evolution, *drift* and *volatility*. These terms are equivalent to the drift and diffusivity, respectively, of the corresponding stochastic process. The drift is the rate that the average of the security price changes, while the volatility is the standard deviation of the return of the asset. There are various techniques, such as *minimum entropy* and the *method of matching moments*, that have been suggested to fit stochastic processes to financial securities.

3.2.1 Discounted price process

A key idea in stochastic modelling of option prices is the *discounted price process*. We have seen that bonds define a risk-free interest rate in our markets, and stated that holding a stock is advantageous only if we expect it to outperform the guaranteed return of a bond. Thus, we should scale the price process of any investment strategy to be discounted by the risk-free interest rate.

Suppose that the risk-free interest rate is $r \geq 0$. Given some stochastic process $\{Y_t \colon t \in \mathbb{R}_+\}$ which may be the price process of a security, derivative, or trading strategy, the discounted price process is another stochastic process $\{Y_t^* \colon t \in \mathbb{R}_+\}$ where $Y^* = Y_t^* = e^{-rt}Y_t$ for $t \in \mathbb{R}_+$. If Y^* has a positive expected return, then the security (or similar) is expected to outperform the bond.

3.2.2 Equivalent martingale measure

We previously discussed a heuristic definition of an *equivalent martingale measure* (EMM). The EMM can now be defined more precisely.

Definition 35 (Equivalent martingale measure)**.** *Suppose that a financial market can be represented by the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and a security on the market is modelling by a stochastic process* $S = \{S_t \colon t \in \mathbb{R}_+\}$*. Then, an* ***equivalent martingale measure***

for $\mathbb{P}$ is another probability measure, $\mathbb{Q}$, such that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ and $\{S_t^* : t \in \mathbb{R}_+\}$ is a martingale under $\mathbb{Q}$.

Risk in holding securities should be considered when calculating a reasonable price. Since investors are (generally) risk-averse, additional profit is demanded to compensate for the risk. This means that the price of a security is less than the expected value of its payout function. However, under an EMM, the security price is exactly equal to the expected value of the discounted process. Accordingly, we usually need to define an EMM in option pricing models.

3.3 Black-Scholes formula

The Black-Scholes formula is an explicit formula for the price of a European call option. It was the first widely-used explicit derivative pricing model and was published by Fischer Black and Myron Scholes in 1973 [5]. The results were based on a stochastic model for the price of a European call option, where the underlying security was assumed to evolve according to a Brownian motion, as was earlier theorised by Louis Bachelier in 1900 [2]. Black and Scholes found a unique equivalent martingale measure and showed that the fair value of the option is the expected value of the discounted price process under this measure [7], deriving the Black-Scholes formula. Robert Merton wrote a follow-up paper in the same year, with various improvements and generalisations [42]. The model led to a significant economic boom by providing mathematical legitimacy to the derivatives markets and allowing them to be implemented more effectively. Accordingly, Merton and Scholes received the 1997 Nobel Memorial Prize in Economic Sciences for the discovery; Black passed away before this time so did not receive the award. In this section, we will provide a brief derivation of the Black-Scholes formula. This will allow us to discuss the shortcomings of the theorem, which motivated the subsequent literature and our research. We follow the derivations in *Introduction to Mathematics of Finance* by Williams [55] and *Non-Gaussian Merton-Black-Scholes Theory* by Boyarchenko [7].

Consider a European call option. Let $T \in (0, \infty)$ be the maturity of the option; trading takes place on the finite interval $[0, T]$. Assume that there is a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that represents the financial market. On this probability space, define a standard Brownian motion $W = \{W_t : t \in [0, T]\}$. Let $(\mathcal{F}_t) = \{\mathcal{F}_t : t \in [0, T]\}$ be the natural filtration generated by W under $\mathbb{P}$. All random variables in this derivation are defined on $(\Omega, \mathcal{F}_t, (\mathcal{F}_t), \mathbb{P})$.

Consider a derivatives market with a riskless bond, $B = \{B_t : t \in [0, T]\}$, that yields a constant return rate $r \geq 0$, which is the risk-free interest rate of the market. Suppose that a risky stock in the market has price $S = \{S_t : t \in [0, T]\}$, which can be modelled according to a geometric Brownian motion, with spot price $S_0 > 0$, drift $\mu \in \mathbb{R}$, and diffusivity (volatility) $\sigma > 0$. So,

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right), \tag{3.3}$$
$$B_t = e^{rt}, \quad t \in [0,T].$$

Let χ be a trading strategy in this market. Then, the discounted stock price process and discounted value process are

$$S_t^* = \frac{S_t}{B_t} = e^{-rt}S_t = S_0 \exp\left(\left(\mu - r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right), \tag{3.4}$$
$$V_t^*(\chi) = \frac{V_t(\chi)}{B_t} = e^{-rt}V_t(\chi), \quad t \in [0,T],$$

respectively.

The Black-Scholes theory assumes that the market is both efficient and complete. These assumptions are satisfied by introducing conditions on the trading strategy, χ, which can be phrased in several different ways. We may prescribe integrability conditions on χ, or require that the discounted value process is either a martingale or is L^2-bounded under an equivalent martingale measure [55].

By Theorem 6, there is a unique equivalent martingale measure, $\mathbb{P}^*$. It turns out that this measure should remove the drift term, $(\mu - r)$, in the discounted price process, (3.4). We can find such a measure using the Girsanov transformation [48] as follows. Let

$$\theta = \frac{\mu - r}{\sigma}, \tag{3.5}$$
$$\Lambda_t = \exp\left(-\theta W_t - \frac{1}{2}\theta^2 t\right), \quad t \in [0,T].$$

Then, according to [12], $\{\Lambda_t \colon t \in [0,T]\}$ is a positive martingale under $\mathbb{P}$. On the measurable space $(\Omega, \mathcal{F})$ endowed with the filtration $(\mathcal{F}_t)$, we define the probability measure $\mathbb{P}^*$ such that

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \Lambda_t \quad \text{on } \mathcal{F}.$$

Since $\Lambda_t > 0$ for all t, we must have $\mathbb{P}^*(A) = 0$ if and only if $\mathbb{P}(A) = 0$ for all $A \in \mathcal{F}$. Thus, $\mathbb{P}^*$ is equivalent to $\mathbb{P}$, according to Definition 7. Now we only require the discounted stock process to be a martingale under $\mathbb{P}^*$. Let

$$\widetilde{W}_t \coloneqq W_t + \theta t, \quad t \in [0,T].$$

By Girsanov's theorem, $\{\widetilde{W}_t \colon t \in [0,T]\}$ is a standard Brownian motion and is a martingale under $\mathbb{P}^*$. Also, by Equations (3.4) and (3.5),

$$\begin{aligned} S_t^* &= S_0 \exp\left(\left(\mu - r - \frac{1}{2}\sigma^2\right)t + \sigma\left(\widetilde{W}_t - \theta t\right)\right) \\ &= S_0 \exp\left(\left(-\frac{1}{2}\sigma^2\right)t + \sigma\widetilde{W}_t\right), \quad t \in [0,T], \end{aligned} \tag{3.6}$$

and we have removed the drift term. The expression in (3.6) is a martingale with respect to $(\mathcal{F}_t)$ under $\mathbb{P}^*$ (*cf.* [12, Theorem 6.2]). So, as required, we have found a measure $\mathbb{P}^*$ which is equivalent to $\mathbb{P}$ such that the discounted price process is a martingale. Therefore, $\mathbb{P}^*$ is the equivalent martingale measure.

Finally, we can compute the arbitrage-free price process for a European call option using our equivalent martingale measure, $\mathbb{P}^*$. Consider a (vanilla) European call option based on the stock in our market, with strike price K. Let $G = (S_T - K)^+$ be the payout of the option, in accordance with (3.1).

Recalling that the arbitrage-free value of the option at time t, $C = \{C_t \colon t \in [0,T]\}$, is equal to the expectation of the discounted price process under the EMM, we have $\mathbb{P}-a.s.$,

$$C_t^* = \mathbb{E}^{\mathbb{P}^*}\left[G^* \mid \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}^*}\left[(S_t^* - K^*)^+ \mid \mathcal{F}_t\right], \quad t \in [0,T], \tag{3.7}$$

where * denotes the discounted process in all cases, which is the original process multiplied by e^{-rt}. Using the expression for S_t^* in (3.6), for each $t \in [0,T]$,

$$S_T^* = S_t^* \frac{S_T^*}{S_t^*} = S_t^* \frac{S_0 \exp\left(\sigma\widetilde{W}_T - \frac{1}{2}\sigma^2 T\right)}{S_0 \exp\left(\sigma\widetilde{W}_t - \frac{1}{2}\sigma^2 t\right)} = S_t^* \exp\left(\sigma(\widetilde{W}_T - \widetilde{W}_t) - \frac{1}{2}\sigma^2(T-t)\right).$$

Substituting this expression into (3.7), then using the fact that S_t^* is $\mathcal{F}_t$-measurable and $\widetilde{W}_T - \widetilde{W}_t$ is independent and normally distributed with mean 0 and variance $T-t$, we can evaluate the expected value in (3.7):

$$\begin{aligned} C_t^* &= \frac{1}{\sqrt{2\pi(T-t)}} \int_{-\infty}^{\infty} \left(S_t^* e^{\sigma y - \frac{1}{2}\sigma^2(T-t)} - K^*\right)^+ e^{-\frac{y^2}{2(T-t)}}\, dy, \quad t \in [0,T], \\ C_T^* &= (S_T^* - K^*)^+. \end{aligned}$$

Transforming our variable such that $z = y/\sqrt{T-t}$ gives us the density function of the standard normal distribution:

$$\begin{aligned} C_t^* &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(S_t^* e^{\sigma z\sqrt{T-t} - \frac{1}{2}\sigma^2(T-t)} - K^*\right)^+ e^{-\frac{z^2}{2}}\, dz, \quad t \in [0,T], \\ C_T^* &= (S_T^* - K^*)^+. \end{aligned} \tag{3.8}$$

Let $\Phi\colon \mathbb{R} \to [0,1]$ be the cumulative distribution function for the standard normal random variable:

$$\Phi(z) = \mathbb{P}(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}}\, du.$$

Then, by manipulating the integral in (3.8) (see [55, Section 4.6.1] for details), we can find the price process of our European call option in terms of the spot price, strike price, volatility, and maturity:

$$C_0 = S_0\Phi\left(\frac{\log\left(\frac{S_0}{K^*}\right)}{\sigma\sqrt{T}} + \frac{1}{2}\sigma^2\sqrt{T}\right) - K^*\Phi\left(\frac{\log\left(\frac{S_0}{K^*}\right)}{\sigma\sqrt{T}} - \frac{1}{2}\sigma^2\sqrt{T}\right).$$

This is the Black-Scholes formula for the price of a European call option. This formula can be extended in a straightforward way to consider European put options, and with slightly more difficulty, American options.

Weaknesses of Black-Scholes

Despite initial success and generally accurate results, the Black-Scholes formula has been shown to have several significant weaknesses. Brownian motion does not have the correct distribution for most securities and does not account for large, sudden jumps in price. Additionally, the assumption of a complete market in the Black-Scholes theory has been shown to be unreasonable in most markets.

The Black-Scholes model assumes that the log-price of a given security can be modelled according to Brownian motion, which has increments that are normally distributed. However, empirical studies of financial markets have provided strong evidence that the log-densities should have fat tails, meaning that the *skewness* and *kurtosis* should be greater than a normal distribution. This is explained in detail and demonstrated through [8, Figure 1.1]. Using a distribution with thinner tails means that more extreme changes in security price have a lower probability of occurring, often leading to the asset being undervalued [8, 38]. Accordingly, the assumption of Brownian motion in the Black-Scholes formula is not reasonable and is a major weakness of the model; other distributions, such as exponential, would be more appropriate.

Additionally, financial assets often exhibit large, sudden price changes, which are not well modelled by continuous processes, such as Brownian motion in the Black-Scholes model. These jumps are caused by sudden events such as the default of a company, an operational failure, a large insurance claim, or a significant information release such as an earnings report [47]. To more accurately model these events, the recent literature has aimed to incorporate jump processes (potentially with drift and diffusions) [8, 38, 47]. These processes have been shown to improve the accuracy of the model, meaning that the continuity of Brownian motion is another weakness of the Black-Scholes formula.

Finally, the Black-Scholes model's assumption that financial markets are complete is not reasonable, in general. Under the assumption of a complete market, there is a replicating strategy for every derivative. This means that options are actually redundant, which is obviously not a desirable property and has been shown to be false in most cases. Alternatively, if the driving Lévy process is neither a Brownian motion nor a Poisson process, the market is incomplete and resultant models would be more reasonable [8]. Thus, the assumption of a complete market is another weakness of the original Black-Scholes model.

Overall, while it contributed significantly to the field of mathematical finance, the Black-Scholes model has several key weaknesses. The literature is attempting to improve these issues by producing models that are driven by more generalised Lévy processes instead of Brownian motion.

Chapter 4

Option pricing models driven by Lévy processes

Over the last two decades, financial mathematics literature has placed a strong emphasis on models driven by Lévy processes. The literature can be categorised into three groups: Monte Carlo simulation (see, *e.g.*, [31, 37, 56]), backward induction (see, *e.g.*, [22, 23, 27, 50]), and integro-differential equations (see, *e.g.*, [7, 10, 29, 30, 32, 43, 45]). These new models aim to remove the weaknesses of the original Black-Scholes model that were identified in Section 3.3. Lévy processes allow us to choose increment distributions that are more appropriate and they admit jumps in their path, which we identified to be a desirable property. The first Lévy-based model was proposed by Mandelbrot in 1963, using a class called *α-stable Lévy processes* to model cotton prices [41]. This area has become much more prevalent in recent times.

Lévy-driven option pricing models assume the underlying asset can be represented by an exponential Lévy process, such that the value of the asset $\{S_t \colon t \in \mathbb{R}_+\}$ is given by

$$S_t = S_0 e^{X_t},$$

where $\{X_t \colon t \in \mathbb{R}_+\}$ is a Lévy process. As we saw in Section 2.3.2, Lévy processes are an extremely rich class with a wide range of properties. Accordingly, there are many choices in the option pricing model, which have been gradually investigated in the literature. It is likely that different Lévy processes will be suited to particular financial markets, so it is beneficial for us to have a wide range of different models.

This chapter discusses the advantages and weaknesses of Lévy-driven models and investigates some relevant techniques. An option pricing algorithm from the recent literature called the *Simple Wiener-Hopf method* is also presented [33].

4.1 Types of Lévy processes used in finance

A variety of families of Lévy processes have been used to model stock returns and option prices in the literature. The most common are *variance gamma processes* (VGP), *normal inverse Gaussian processes* (NIG), *hyperbolic processes* (HP), *KoBoL processes*, and *normal tempered stable Lévy processes* [8]. The probability densities of these families have semi-heavy (exponentially decaying) tails, which have been shown to more accurately model financial securities. The Lévy measures of these processes also have polynomial singularities at the origin. Processes that decay exponentially and have a polynomial singularity at the origin are known as *regular Lévy processes of exponential type* (RLPEs). The Lévy processes used in financial modelling also often allow jumps in their path, which more effectively captures the large, sudden changes in security prices that often occur. Here, we mainly focus on KoBoL processes to demonstrate existing methods.

4.1.1 KoBoL processes

KoBoL processes are an extension of another class, truncated Lévy processes, and were introduced by Boyarchenko and Levendorskiĭ in 1999 [9]. In this paper, and various others since, it has been shown that the KoBoL family provides a good fit for many financial securities. KoBoL processes evolve according to a pure jump process with drift and are defined in terms of their Lévy measure as follows.

Definition 36 (KoBoL family)**.** *A Lévy process, X, belongs to the **KoBoL family** if it is purely discontinuous ($\sigma = 0$) and its Lévy measure has the form*

$$\Pi(dx) = c_+\Pi^+(\nu, -\lambda_-; dx) + c_-\Pi^-(\nu, \lambda_+; dx),$$

where $c_\pm > 0$, $\lambda_- < 0 < \lambda_+$, and

$$\Pi^+(\nu, \lambda; dx) = x_+^{-\nu-1}e^{-\lambda x}\,dx, \quad \Pi^-(\nu, \lambda; dx) = x_-^{-\nu-1}e^{\lambda x}\,dx,$$

*where $\nu < 2$, $\lambda > 0$, $x_+ = \max\{x, 0\}$, and $x_- = x_+ - x$. The parameter ν is called the **order** of the KoBoL process.*

The characteristic exponent of a KoBoL process can be easily calculated by substituting the Lévy measure into Definition 22 [8, Section 3.1.2.2].

Lemma 7. *Let X be a KoBoL process of order ν with characteristic exponent $\Psi(\xi)$. Then,*

i. If $\nu < 2$, $\nu \neq 0, 1$,

$$\Psi(\xi) = -i\mu\xi + c_+\Gamma(-\nu)[(-\lambda_-)^\nu - (-\lambda_- - i\xi)^\nu] + c_-\Gamma(-\nu)\big[\lambda_+^\nu - (\lambda_+ + i\xi)^\nu\big].$$

ii. *If* $\nu = 0$,

$$\Psi(\xi) = -i\mu\xi + c_+[\ln(-\lambda_- - i\xi) - \ln(-\lambda_-)] + c_-[\ln(\lambda_+ + i\xi) - \ln(\lambda_+)].$$

iii. *If* $\nu = 1$,

$$\begin{aligned}\Psi(\xi) = -i\mu\xi &+ c_+[(-\lambda_-)\ln(-\lambda_-) - (-\lambda_- - i\xi)\ln(-\lambda_- - i\xi)] \\ &+ c_-[\lambda_+ \ln(\lambda_+) - (\lambda_+ + i\xi)\ln(\lambda_+ + i\xi)].\end{aligned}$$

The KoBoL process has appeared regularly in financial modelling literature (see, *e.g.*, [9, 32]), including in the Simple Wiener-Hopf method discussed in Section 4.4.

4.2 Equivalent martingale measures in Lévy models

Choosing an appropriate equivalent martingale measure (EMM) is a key step in producing an option pricing model, as discussed in Section 3.2.2. According to Theorem 6, there is exactly one EMM in a complete market. The Black-Scholes theory made this assumption, making the choice of measure straightforward. However, real markets are rarely complete; recent Lévy models aim to derive option prices in incomplete markets. We still assume market efficiency, so, by Theorem 5, there is at least one EMM, but we need to choose the appropriate measure. There are several methods that have been suggested for this task, including minimising entropy of the measures or various other optimisation arguments discussed in [8, Chapter 4]. Commonly, the *Esscher transform* is used, which we will briefly present in Section 4.2.1.

The choice of the change of measure depends on the underlying Lévy process and empirical studies of the financial market that is under consideration. The results in the following sections and chapters require that an EMM has been chosen, but do not depend on the particular choice. For any EMM, $\mathbb{Q}$, the underlying Lévy process must satisfy the EMM condition [20, 33]. This formula involves the *dividend yield*, $d \geq 0$, of the asset, which we will always assume to be 0; alterations to results are straightforward if $d > 0$.

Lemma 8 (EMM condition)**.** *Consider a financial security that is modelled by an exponential of some Lévy process,* $X = \{X_t \colon t \in \mathbb{R}_+\}$*, with dividend yield* $d \geq 0$ *and risk-free interest rate* $r \geq 0$*. The EMM condition states that*

$$\mathbb{E}^{\mathbb{Q}}\big[e^{X_t}\big] = e^{(r-d)t}, \tag{4.1}$$

where $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,]$ *is the expectation with respect to the EMM,* $\mathbb{Q}$*.*

By Definition 22, the left-hand side of the EMM condition (4.1) has the following relationship with the characteristic exponent of the Lévy process (defined on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$) evaluated at $-i$, $\Psi(-i)$:

$$\mathbb{E}^{\mathbb{Q}}\big[e^{X_t}\big] = \mathbb{E}^{\mathbb{Q}}\big[e^{i(-i)X_t}\big] = e^{-t\Psi(-i)}.$$

Thus, the EMM condition is equivalent to $e^{-t\Psi(-i)} = e^{(r-d)t}$, which implies that

$$r - d + \Psi(-i) = 0.$$

Finally, evaluating $\Psi(-i)$ according to the Lévy-Khintchine formula in Theorem 2, we derive a relationship between the drift, μ, and the other parameters of the process that is equivalent to the EMM condition:

$$\mu = r - d + \frac{\sigma^2}{2} + \int_{\mathbb{R}} \left(1 - e^x + x\mathbb{1}\{|x| < 1\}\right) \Pi(dx). \tag{4.2}$$

Under an EMM, the Lévy process that drives our model must satisfy this condition. Accordingly, we can use (4.2) to calculate the prescribed drift for our process under the risk-neutral measure.

4.2.1 Esscher transform

Let $X = \{X_t \colon t \in \mathbb{R}_+\}$ be a Lévy process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. The Esscher transform of $\mathbb{P}$ is another measure, $\mathbb{Q}$, which is equivalent to $\mathbb{P}$ and satisfies

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{F_t} = e^{\theta X_t - d(\theta,t)}, \tag{4.3}$$

where $d(\theta, t)$ is a constant. In [8, Section 4.1.1], the authors demonstrate that the Esscher transform usually exists for the processes used in financial studies.

4.3 Option pricing models

Consider a knock-out put barrier option with strike price K, barrier H, spot price S_0, and maturity T. Suppose that the underlying security is modelled by an exponential Lévy process, $S_t = S_0 e^{X_t}$, for some Lévy process $X = \{X_t \colon t \in \mathbb{R}_+\}$ on the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$. Choose some EMM, $\mathbb{Q}$, and define X on the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{Q})$ such that X now satisfies the EMM condition and (4.2). All of the expectations in this chapter are taken with respect to this EMM.

Let $h \coloneqq \ln\left(\frac{H}{S_0}\right)$ be the value of the barrier when compared to the log-price, X. Let $V(T, t, x)$ be the no-arbitrage value of the option with maturity T, at time t, when the underlying Lévy process has value x and the security has value $S_0 e^x$. Despite initially being a fixed parameter, the maturity is included as a variable in the function V because it will be manipulated later.

The evaluation of the current price of the option depends on whether the process has reached the barrier or maturity. Let $E = (-\infty, T] \times \mathbb{R}$ be an auxiliary sample space; on E, define $\widehat{X} = \{\widehat{X}_t = (t, X_t) \colon t \in \mathbb{R}_+\}$. Let $C = (-\infty, T) \times (h, \infty)$ be the region where $\widehat{X}$

is above the barrier and before maturity, where the option has positive value and is still being observed. Let $B = E \setminus C$ be the region where we would stop observing the option, either because it has reached maturity or fallen below the barrier. The two regions, B and C, are shown in Figure 4.1, for the same example barrier option from Figure 3.3.

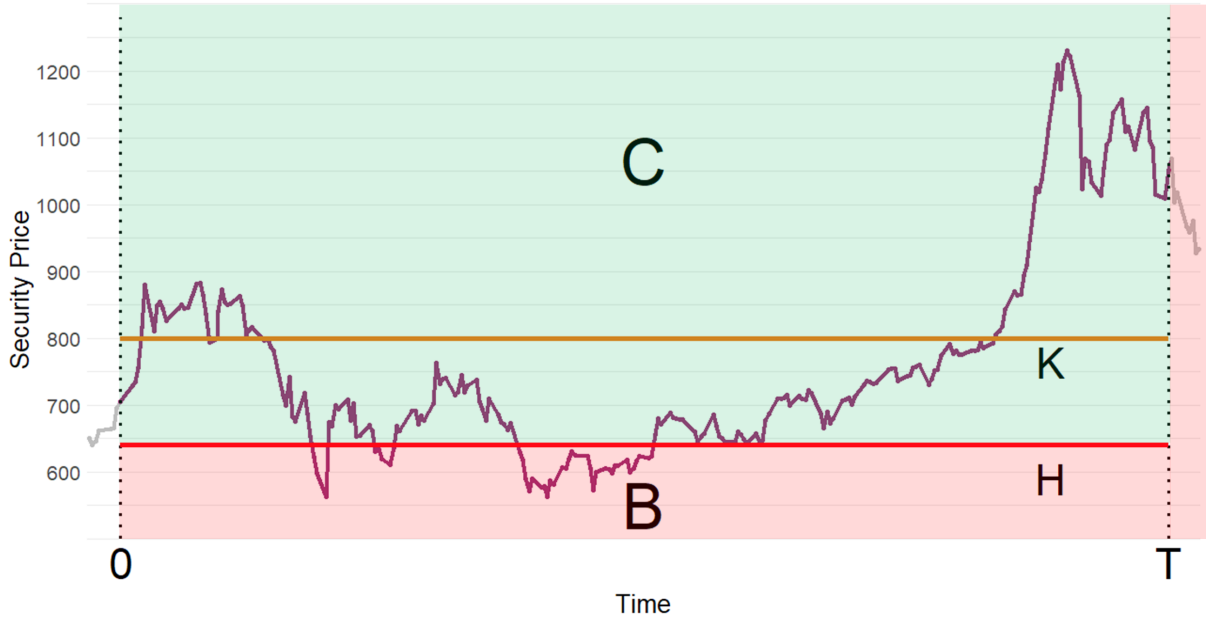


Figure 4.1: Barrier option depicting regions B and C in red and green, respectively.

For $\widehat{X}_t \in B$, the value of the option is immediately given by

$$G(X_t, \underline{X}_t) := (K - e^{X_t})^+ \mathbb{1}\{\underline{X}_t > h\}, \tag{4.4}$$

which represents the payout function of the barrier option. Alternatively, the option value in region C is determined by its expected value when it hits region B. Accordingly, let $\tau_B = \{t \in \mathbb{R}_+ : \overline{X}_t \in B\}$ be the hitting time of X to the region B. By definition, under a chosen EMM, the option price is a (local) martingale. Hence, since $\tau_B \leq T < \infty$, by the optional stopping theorem, the no-arbitrage value of the option at time $t \in [0, \tau_B]$ and current process value $x \in C$ must satisfy

$$V(T, t, x) = \mathbb{E}^x\left[e^{-r(\tau_B - t)} G(X_{\tau_B}, \underline{X}_{\tau_B})\right], \tag{4.5}$$

where $\mathbb{E}^x[\cdot] = \mathbb{E}[\cdot \mid X_0 = x]$. We are only interested in the value of the option at the start of the contract, when $t = 0$, which is given by

$$V(T, x) := V(T, 0, x) = \mathbb{E}^x\left[e^{-r\tau_B} G(X_{\tau_B}, \underline{X}_{\tau_B})\right]. \tag{4.6}$$

There are several methods to evaluate (4.6) and find the option value. First, we may calculate the *Laplace transform* of this expectation and then apply the *Post-Widder inversion formula*. Alternatively, an equivalent boundary problem can be derived and then solved using a method called *Carr's randomisation*, which we review later.

4.3.1 Laplace transform method

First, the hitting time to region B, τ_B, occurs when we either reach the barrier or maturity. The effect of hitting the barrier is already captured by the payout function in (4.4). Thus, τ_B in (4.6) could be replaced by the hitting time to the maturity date, which is deterministic with value T. This means that the option pricing formula in (4.6) is equivalent to the following equation.

$$V(T,x) = \mathbb{E}^x\big[e^{-rT}G(X_T, \underline{X}_T)\big]. \tag{4.7}$$

A solution for this equation can be derived by first taking the Laplace transform, then using the Wiener-Hopf factorisation, *Fourier transform*, and *expected present value (EPV) operators*. The ideas for this derivation were given in [30].

Preliminaries

The Laplace transform converts a real-valued function into a complex-valued function in the *frequency domain* and is often used to solve differential equations.

Definition 37 (Laplace transform)**.** *Given a function $f : \mathbb{R}_+ \to \mathbb{R}$, the **Laplace transform** of f is*

$$\mathcal{L}\{f\}(q) = \int_0^\infty e^{-qt} f(t)\, dt = \mathbb{E}\big[e^{-qX}\big], \quad \text{for } q \in \mathbb{C},$$

where X is a random variable with probability density function f.

Various techniques to invert the Laplace transformed function into the correct domain are available. We will use the Post-Widder formula, which was introduced and proved in [54, Section 8.2].

Theorem 7 (Post-Widder formula)**.** *Given the Laplace transform, $\widehat{f}$, of a function $f : \mathbb{R} \to \mathbb{C}$, define for each $p \in \mathbb{N}$,*

$$f_p(t) = \frac{(-1)^p}{p!}\Big(\frac{p}{t}\Big)^{p+1} \widehat{f}^{(p)}\Big(\frac{p}{t}\Big), \tag{4.8}$$

*where $\widehat{f}^{(p)}(t)$ is the p-th derivative of the $\widehat{f}(t)$. Then, the **Post-Widder formula** states that*

$$f(t) = \lim_{p\to\infty} f_p(t).$$

This procedure can also be applied to a multivariable function, where the Laplace transform is taken with respect to one particular variable. The Post-Widder formula (4.8) has the same form, except with additional arguments of the function; only the subject of the Laplace transform is altered. Thus, for a function of two variables, $f(t, x)$, the formula that is equivalent to (4.8) would be

$$f_p(t,x) = \frac{(-1)^p}{p!}\left(\frac{p}{t}\right)^{p+1} \widehat{f}^{(p)}\left(\frac{p}{t},\, x\right).$$

The Post-Widder formula converges relatively slowly; the convergence is $\mathcal{O}(p^{-1})$ and usually requires about 500 steps for sufficient accuracy [30]. An acceleration technique was introduced in [1], where a linear combination of the Post-Widder terms is used to reduce the total number of calculations.

Theorem 8 (Post-Widder acceleration)**.** *Given $f_p(x)$ from the Post-Widder formula, define*

$$f_{p,m}(x) = \sum_{k=1}^{m} w(k,m) f_{pk}(x),$$

where

$$w(k,m) = (-1)^{m-k} \frac{k^m}{k!(m-k)!}.$$

Then,

$$f(x) = \lim_{p,m\to\infty} f_{p,m}(x).$$

The accelerated Post-Widder formula has convergence rate $\mathcal{O}(p^{-m})$, and parameter values of approximately $p = 5$ and $m = 3$ are usually sufficient [30]. The computational complexity is significantly reduced and speed is increased.

The *Fourier transform* is a particular case of the Laplace transform where the complex argument, q, is replaced by a strictly imaginary variable, $i\xi$.

Definition 38 (Fourier transform)**.** *The Fourier transform of a function $f\colon \mathbb{R} \to \mathbb{C}$ is*

$$\mathcal{F}\{f\}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} f(x)\, dx, \quad \textit{for } \xi \in \mathbb{R}.$$

The inverse Fourier transform is

$$\mathcal{F}^{-1}\{\widehat{f}\}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \widehat{f}(\xi)\, d\xi, \quad \textit{for } x \in \mathbb{R}.$$

Fourier transforms are often used to simplify differential equations and expectations. For example, a derivative becomes multiplication by a variable after a Fourier transform, as observed by considering the transform of the derivative, $f'(x)$, of some function, $f(x)$. Indeed, integrating by parts,

$$\begin{aligned}\mathcal{F}\{f'\}(\xi) &= \int_{\mathbb{R}} e^{-ix\xi} f'(x)dx \\ &= \left[f(x)e^{-ix\xi}\right]_{-\infty}^{\infty} - \int_{\mathbb{R}} f(x)\frac{d}{dx}\left[e^{-ix\xi}\right] dx \\ &= i\xi \int_{\mathbb{R}} f(x)e^{-ix\xi}\, dx \\ &= i\xi \widehat{f}(\xi).\end{aligned}$$

The simplified problem can be solved and then transformed back to the original domain by the inverse Fourier transform. We use this technique, by defining EPV operators which apply Fourier transform inversion to the Wiener-Hopf factors.

The expected present value operator, also known as the *normalised resolvent*, represents the Wiener-Hopf factorisation in operator form [7].

Definition 39 (Expected present value operator)**.** *For a Lévy process $\{X_t : t \in \mathbb{R}_+\}$ and a function $u : \mathbb{R} \to \mathbb{C}$, the expected present value operator of u with parameter q is*

$$\mathcal{E}_q u(x) \coloneqq \mathbb{E}\left[u\left(x + X_{e_q}\right)\right].$$

The EPV operator is named as such because, due to the stationary increments property of Lévy processes, it can equivalently be written as

$$\mathcal{E}_q u(x) = \mathbb{E}^x\left[\int_{\mathbb{R}_+} qe^{-qt}u(X_t)\, dt\right],$$

which is the expected value of u after a random exponential amount of time.

To represent the Wiener-Hopf factorisation through these operators, we define an EPV operator under the driving Lévy process, along with its supremum and infimum processes:

$$\mathcal{E}_q u(x) \coloneqq \mathbb{E}\left[u\left(x + X_{e_q}\right)\right], \tag{4.9}$$

$$\mathcal{E}_q^+ u(x) \coloneqq \mathbb{E}\left[u(x + \overline{X}_{e_q})\right], \tag{4.10}$$

$$\mathcal{E}_q^- u(x) \coloneqq \mathbb{E}\left[u(x + \underline{X}_{e_q})\right]. \tag{4.11}$$

Then, the Wiener-Hopf factorisation can be written in operator form as

$$\mathcal{E}_q = \mathcal{E}_q^+ \mathcal{E}_q^- = \mathcal{E}_q^- \mathcal{E}_q^+.$$

To numerically implement the EPV operators, we derive a formula in terms of the Fourier transform and inverse Fourier transform, using the following lemma.

Lemma 9. *Let $u\colon \mathbb{R} \to \mathbb{R}_+$ be a non-negative function and Y be a random variable with characteristic function ϕ and probability density f_Y. Then,*

$$\mathbb{E}[u(x+Y)] = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi}\, \phi(\xi)\, \widehat{u}(\xi)\, d\xi, \quad x \in \mathbb{R},$$

where $\widehat{u}$ is the Fourier transform of u.

Proof. This relationship is derived by applying the relevant definitions along with Tonelli's theorem:

$$\begin{aligned}
&\frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi}\, \phi(\xi)\, \widehat{u}(\xi)\, d\xi \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi}\, \phi(\xi) \int_{\mathbb{R}} e^{-iy\xi}\, u(y)\, dy\, d\xi && \text{(applying Definition 38)} \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-i\xi(y-x)}\, \phi(\xi)\, u(y)\, d\xi\, dy && \text{(by Tonelli's theorem)} \\
&= \int_{\mathbb{R}} u(y)\, \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi(y-x)}\, \phi(\xi)\, d\xi\, dy \\
&= \int_{\mathbb{R}} u(y)\, f_Y(y-x)\, dy && \text{(applying the inversion formula (2.9))} \\
&= \int_{\mathbb{R}} u(x+z)\, f_Y(z)\, dz \\
&= \mathbb{E}[u(x+Y)].
\end{aligned}$$

□

Finally, let $\phi_q(\xi)$, $\phi_q^+(\xi)$, and $\phi_q^-(\xi)$ be the characteristic functions of the random variables $X_{\mathbf{e}_q}$, $\overline{X}_{\mathbf{e}_q}$, and $\underline{X}_{\mathbf{e}_q}$, respectively. Then, using Lemma 9,

$$\begin{aligned}
\mathcal{E}_q^+ u(x) &= \mathbb{E}\big[u(x + \overline{X}_{e_q})\big] \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi}\, \phi^+(\xi)\, \widehat{u}(\xi)\, d\xi \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi}\, \phi^+(\xi) \int_{\mathbb{R}} e^{-ix\xi}\, u(x)\, dx\, d\xi \\
&= \mathcal{F}^{-1}\{\, \phi_q^+(\xi)\, \mathcal{F}\{u\}(\xi)\}(x), \qquad (4.12)
\end{aligned}$$

and similarly,

$$\mathcal{E}_q^- u(x) = \mathcal{F}^{-1}\{\, \phi_q^-(\xi)\, \mathcal{F}\{u\}(\xi)\}(x). \qquad (4.13)$$

The brackets in these formulae are often distracting and will thus be excluded where appropriate. We now have an operator form of the Wiener-Hopf factorisation that can

be calculated using the Fourier transform, which is efficiently implemented numerically using the *fast Fourier transform* discussed in Appendix A.2.

All of the required techniques for the Laplace transform method of calculating the option price have been defined and we now present the model derivation.

Pricing method

Now we apply these techniques to solve the option pricing equation in (4.7),

$$V(T,x) = \mathbb{E}^x\big[e^{-rT}G(X_T, \underline{X}_T)\big].$$

The Laplace transform of this function with respect to the maturity T is

$$\widehat{V}(q,x) = \int_{\mathbb{R}_+} e^{-qt}\, V(t,x)\, dt = \int_{\mathbb{R}_+} e^{-(q+r)t}\, \mathbb{E}^x[G(X_t, \underline{X}_t)]\, dt.$$

The new variable, q, is called the *observation rate* because it defines the time between points in our numerical algorithm. To implement the Post-Widder formula and calculate the option price, $V(T,x)$, we need to calculate the p-th derivative of the Laplace transform. For $p = 0, 1, 2, \ldots$, using the Leibniz integration rule,

$$\begin{aligned}
\widehat{V}^{\,(p)}(q,x) &= \int_{\mathbb{R}_+} (-t)^p\, e^{-(q+r)t}\, \mathbb{E}^x[G(X_t, \underline{X}_t)]\, dt \\
&= (-1)^p\, \frac{\Gamma(p+1)}{(q+r)^{p+1}} \int_{\mathbb{R}_+} \frac{(q+r)^{p+1}}{\Gamma(p+1)}\, t^p\, e^{-(q+r)t}\, \mathbb{E}^x[G(X_t, \underline{X}_t)]\, dt \\
&= (-1)^p\, \frac{\Gamma(p+1)}{(q+r)^{p+1}}\, \mathbb{E}^x\big[G(X_{\mathbf{g}(p+1,\,q+r)}, \underline{X}_{\mathbf{g}(p+1,\,q+r)})\big],
\end{aligned}$$

where $\Gamma(z)$ is the gamma function, and $\mathbf{g}(p+1, q+r)$ is an independent Gamma-distributed random variable with shape parameter $p+1 > 0$ and rate $q+r > 0$, which has probability density function $f(t) = \frac{(q+r)^{p+1}}{\Gamma(p+1)}\, t^p\, e^{-(q+r)t}$. Since, for $p \in \mathbb{N}$, $\Gamma(p+1) = p!$, we have

$$\widehat{V}^{\,(p)}(q,x) = (-1)^p\, \frac{p!}{(q+r)^{p+1}}\, \mathbb{E}^x\big[G(X_{\mathbf{g}(p+1,\,q+r)}, \underline{X}_{\mathbf{g}(p+1,\,q+r)})\big].$$

Finally, setting $q = \frac{p}{T}$ and multiplying by the coefficient given by the Post-Widder formula, for each $p = 0, 1, 2, \ldots$, define

$$\begin{aligned}
v_p(T,x) &\coloneqq \frac{(-1)^p}{p!} \Big(\frac{p}{T}\Big)^{p+1} \widehat{V}^{\,(p)}\Big(\frac{p}{T},\, x\Big) \\
&= \frac{(-1)^p}{p!}\, q^{p+1}\, \widehat{V}^{\,(p)}(q,\, x) \\
&= \left(\frac{q}{q+r}\right)^{p+1} \mathbb{E}^x\big[G(X_{\mathbf{g}(p+1,\,q+r)}, \underline{X}_{\mathbf{g}(p+1,\,q+r)})\big]. \qquad (4.14)
\end{aligned}$$

By the Post-Widder formula, $v_p(T,x)$ converges to the option price, $V(T,x)$, as p approaches infinity. Provided that $\mathbb{E}^x\big[G(X_{\mathbf{g}(p+1,\,q+r)}, \underline{X}_{\mathbf{g}(p+1,\,q+r)})\big]$ can be calculated, we have our option pricing formula. This final step of the derivation depends on the form of the payout function, $G(x,y)$. Recall from (4.4),

$$G(X_t, \underline{X}_t) = (K - e^{X_t})^+ \mathbb{1}\{\underline{X}_t > h\}.$$

Letting $\widetilde{G}(x) = (K - e^x)^+$, we can write

$$G(X_t, \underline{X}_t) = \widetilde{G}(X_t)\, \mathbb{1}\{\underline{X}_t > h\},$$

so that

$$v_p(T,x) = \left(\frac{q}{q+r}\right)^{p+1} \mathbb{E}^x\Big[\widetilde{G}\big(X_{\mathbf{g}(p+1,\,q+r)}\big)\, \mathbb{1}\{\underline{X}_{\mathbf{g}(p+1,\,q+r)} > h\}\Big]. \tag{4.15}$$

Both of the terms in the expectation in (4.15) involve a Gamma random variable with integer shape parameter $p+1 \in \mathbb{Z}$. When the shape parameter is an integer, the Gamma distribution is equivalent to the Erlang distribution with the same parameters, which is defined as the sum of $(p+1)$ independent exponential variables with rate $(q+r)$. One term from this independent sum can be separated out to give the following relationship.

$$\mathbf{g}(p+1,\, q+r) \overset{d}{=} \mathbf{g}(p,\, q+r) + \mathbf{e}_{q+r}, \tag{4.16}$$

where $\mathbf{e}_{q+r} \sim \text{Exp}(q+r)$, independent of $\mathbf{g}(p, q+r)$. Without loss of generality, the equality in (4.16) can be taken almost surely instead of in distribution. This leads to a similar decomposition for the indicator function from (4.15), where we split the Gamma variable into the sum of Gamma and exponential variables. Due to the independent increment property of X, the infimum of the process should be greater than h at both the exponential and Gamma realisations. Given this, the sum of the process at the exponential time and the infimum of the process at the Gamma time should also be greater than h. Thus, we have the following relationship.

$$\mathbb{1}\{\underline{X}_{\mathbf{g}(p+1,\,q+r)} > h\} = \mathbb{1}\{X_{\mathbf{e}_{q+r}} + \underline{X}_{\mathbf{e}_q \le\, s\, \le \mathbf{e}_q + \mathbf{g}(p,\,q+r)} > h\}\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}. \tag{4.17}$$

Applying (4.16) and (4.17) to the formula for v_p in (4.15), gives the following expression.

$$v_p(T,x) = \left(\frac{q}{q+r}\right)^{p+1} \mathbb{E}^x\Big[\widetilde{G}\big(X_{\mathbf{g}(p,\,q+r)+\mathbf{e}_{q+r}}\big)\, \mathbb{1}\{X_{\mathbf{e}_{q+r}} + \underline{X}_{\mathbf{e}_q \le\, s\, \le \mathbf{e}_q + \mathbf{g}(p,\,q+r)} > h\}\, \mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big].$$

Then, by the law of total expectation (tower rule) and the stationary increments property of X,

$$v_p(T,x) = \frac{q}{q+r}\, \mathbb{E}^x\left[\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\left(\frac{q}{q+r}\right)^{p} \mathbb{E}^{X_{\mathbf{e}_{q+r}}}\left[\widetilde{G}(X_{\mathbf{g}(p,\,q+r)})\mathbb{1}\{\underline{X}_{\mathbf{g}(p,\,q+r)} > h\}\right]\right],$$

where $\mathbb{E}^{X_{\mathbf{e}_{q+r}}}[\,\cdot\,]$ is the expectation conditioned on the starting point $X_{\mathbf{e}_{q+r}}$. Finally, recalling the definition of v_p in Equation (4.15), observe that

$$\left(\frac{q}{q+r}\right)^p \mathbb{E}^{X_{\mathbf{e}_{q+r}}}\Big[\widetilde{G}(X_{\mathbf{g}(p,\,q+r)})\,\mathbb{1}\{\underline{X}_{\mathbf{g}(p,\,q+r)} > h\}\Big] =: v_{p-1}\big(T, X_{e_{q+r}}\big).$$

Then, the previous equation gives us the following iterative formula to calculate $v_p(T,x)$ for $p = 0, 1, 2, \ldots$:

$$v_p(T,x) = \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}\big(T,\, X_{\mathbf{e}_{q+r}}\big)\,\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big]. \tag{4.18}$$

Reassign $p \mapsto p+1$ so that (4.18) applies for $p \in \mathbb{N}$. To determine the initial term, consider $v_1(T,x)$ using (4.15) and the fact that $\mathbf{g}(1,\lambda) = \mathbf{e}_\lambda$:

$$v_1(T,x) = \frac{q}{q+r}\,\mathbb{E}^x\Big[\widetilde{G}(X_{\mathbf{g}(1,\,q+r)})\,\mathbb{1}\{X_{\mathbf{g}(1,\,q+r)} > h\}\Big] = \frac{q}{q+r}\,\mathbb{E}^x\Big[\widetilde{G}(X_{\mathbf{e}_{q+r}})\,\mathbb{1}\{X_{\mathbf{e}_{q+r}} > h\}\Big].$$

Defining $v_0(T,x)$ such that $v_1(T,x)$ satisfies (4.18) requires

$$v_0(T,x) = \widetilde{G}(x).$$

Thus, we can now calculate $v_p(T,x)$ for any $p = 0, 1, 2, \ldots$ by iteratively defining

$$\begin{aligned} v_0(T,x) &= \widetilde{G}(x) = (K - e^x)^+, \\ v_p(T,x) &= \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}(T,\, X_{\mathbf{e}_{q+r}})\,\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big]. \end{aligned} \tag{4.19}$$

This expectation can be manipulated using the Wiener-Hopf factorisation in (2.6), the law of total expectation, and then the EPV operators from Definition 39, as follows:

$$\begin{aligned} \mathbb{E}^x\Big[v_{p-1}(T,\, X_{\mathbf{e}_{q+r}})\mathbb{1}\{\underline{X}_{e_{q+r}} > h\}\Big] &= \mathbb{E}^x\Big[v_{p-1}\big(T,\, (X_{\mathbf{e}_{q+r}} - \underline{X}_{\mathbf{e}_{q+r}}) + \underline{X}_{\mathbf{e}_{q+r}}\big)\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big] \\ &= \mathbb{E}^x\Big[\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\, v_{p-1}(T,\, \overline{X}_{\mathbf{e}_{q+r}} + \underline{X}_{\mathbf{e}_{q+r}})\Big] \\ &= \mathbb{E}^x\Big[\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\, \mathbb{E}\big[v_{p-1}(T,\, \overline{X}_{\mathbf{e}_{q+r}} + \underline{X}_{\mathbf{e}_{q+r}})\big]\Big] \\ &= \mathbb{E}^x\Big[\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\, \mathcal{E}^+_{q+r}\, v_{p-1}(T, \underline{X}_{\mathbf{e}_{q+r}})\Big] \\ &= \mathbb{E}\Big[\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\, \mathcal{E}^+_{q}\, v_{p-1}(T, x + \underline{X}_{\mathbf{e}_{q+r}})\Big] \\ &= \mathcal{E}^-_{q+r}\, \mathbb{1}_{(h,\infty)}\, \mathcal{E}^+_{q+r}\, v_{p-1}(T,x). \end{aligned}$$

Finally, using the representation for the EPV operators in (4.12) and (4.13), we now have an iterative formula for $v_p(T,x)$ which is numerically tractable:

$$\begin{aligned} v_0(T,x) &= (K - e^x)^+ \mathbb{1}\{x > h\}, \\ v_p(T,x) &= \frac{q}{q+r}\,\mathcal{F}^{-1}\,\phi^-_{q+r}(\xi)\,\mathcal{F}\,\mathbb{1}_{(h,\infty)}\,\mathcal{F}^{-1}\,\phi^+_{q+r}(\xi)\,\mathcal{F}\,v_{p-1}(T,x), \quad \text{for } p \in \mathbb{N}. \end{aligned} \tag{4.20}$$

Using this formula, we can calculate $v_M(M/T,\, x)$ for large M, which is approximately equal to $V(T,x)$, the option value. This forms our Lévy-driven option pricing algorithm.

4.3.2 Carr's randomisation

Carr's randomisation is another technique that may be used to solve the option pricing formula in (4.6),

$$V(T,x) = \mathbb{E}^x\big[e^{-r\tau_B} G(X_{\tau_B}, \underline{X}_{\tau_B})\big],$$

and model the value of a barrier option. While the final outcome is identical to the Laplace transform method (as we would hope), Carr's randomisation involves vastly different techniques and also presents a probabilistic interpretation. Observe that, since τ_B is the first time that X hits B, we could remove the infimum from the payout function when it is observed at the hitting time. This also means that we no longer need to consider the entire path of the process. Thus, the option pricing formula is equivalent to the following equation:

$$V(T,x) = \mathbb{E}^x\big[e^{-r\tau_B} G_B(X_{\tau_B})\big], \tag{4.21}$$

where

$$G_B(x) = (K - e^x)^+ \mathbb{1}\{x > h\}. \tag{4.22}$$

A boundary *integro-differential equation* can be derived from these formulae and then solved using Carr's randomisation.

Preliminaries

Given a mathematical model, *Randomisation* is a generic procedure with three key steps:

1. Randomise one of the model parameters based on a given distribution with mean equal to the true parameter value.
2. Calculate the expected value of the dependent variable based on the randomised parameter.
3. Let the variance of the governing distribution approach 0.

Randomisation may be used to solve a wide variety of problems, particularly differential equations. Within standard option pricing, we could randomise the initial stock price, the strike price, or the maturity date, as demonstrated in [11]. For a barrier option, the barrier value could also be randomised. In [11], Carr demonstrated that this technique often simplifies the option pricing formula, and thus, it is often referred to as Carr's randomisation in this context.

The main theorem required to derive a boundary problem for the option price is *Dynkin's formula*, which involves the *infinitesimal generator* of our Lévy process. In these statements and subsequent sections, C^k denotes the space of functions that have k continuous derivatives.

Definition 40 (Infinitesimal generator)**.** *Let $X = \{X_t \colon t \in \mathbb{R}_+\}$ be a stochastic process. The **infinitesimal generator** of X is an operator, L, that applies to twice-differentiable functions $f \in C^2$ that vanish at infinity, defined by*

$$(Lf)(x) \coloneqq \lim_{h \downarrow 0} \frac{\mathbb{E}[f(x + X_h)] - f(x)}{h} = \lim_{h \downarrow 0} \frac{\mathbb{E}^x[f(X_h)] - f(x)}{h}, \quad \text{for } x \in \mathbb{R}.$$

Observe that, for small t,

$$\mathbb{E}^x[f(X_t)] \approx f(x) + tLf(x),$$

which shows that the generator describes the movement of the stochastic process in an infinitesimal time interval. The infinitesimal generator of a stochastic process combined with its index time is given by the following lemma.

Lemma 10. *Let $X = \{X_t \colon t \in \mathbb{R}_+\}$ be a stochastic process with infinitesimal generator L. The combined time process, $\widehat{X} = \{\widehat{X}_t = (t, X_t) \colon t \in \mathbb{R}_+\}$, has infinitesimal generator $\widehat{L} = \partial_t + L$, where $\partial_t f = \frac{\partial f}{\partial t}$.*

Proof. Let $f(t, x)$ be any twice-differentiable function that vanishes at infinity. Then, by Definition 40, using the notation $\mathbb{E}^{x,t}[\,\cdot\,] = \mathbb{E}[\,\cdot \mid X_t = x]$,

$$\begin{aligned}
(\widehat{L}f)(t,x) &= \lim_{h \downarrow 0} \frac{\mathbb{E}^{x,t}[f(\widehat{X}_{t+h})] - f(t,x)}{h} \\
&= \lim_{h \downarrow 0} \frac{\mathbb{E}^{x,t}[f(t+h, X_{t+h}) - f(t, X_t)]}{h} \\
&= \lim_{h \downarrow 0} \frac{\mathbb{E}^{x,t}[f(t+h, X_t) - f(t, X_t)] + \mathbb{E}^{x,t}[f(t+h, X_{t+h}) - f(t+h, X_t)]}{h} \\
&= \lim_{h \downarrow 0} \frac{f(t+h, x) - f(t,x)}{h} + \lim_{h \downarrow 0} \frac{\mathbb{E}^{x,t}[f(t+h, X_{t+h})] - f(t+h, x)}{h} \\
&= \lim_{h \downarrow 0} \frac{f(t+h, x) - f(t,x)}{h} + \lim_{h \downarrow 0} \frac{\mathbb{E}^{x}[f(t+h, X_h)] - f(t+h, x)}{h} \\
&= \frac{\partial f}{\partial t} + (Lf)(t,x).
\end{aligned}$$

Thus, $\widehat{L} = \partial_t + L$. □

Now we can state Dynkin's formula, as given in [8, Section 2.3.4]. This can be interpreted as a stochastic generalisation of the second fundamental theorem of calculus.

Theorem 9 (Dynkin's formula)**.** *Consider a stochastic process $X = \{X_t \colon t \in \mathbb{R}_+\}$ with infinitesimal generator L. Let $\lambda > 0$ and $\tau \in \mathbb{R}$ be a stopping time with $\mathbb{E}[\tau] < \infty$. Given a function $f \in C^0$, **Dynkin's formula** states that*

$$f(x) = \mathbb{E}^x\left[e^{-\lambda\tau} f(X_\tau)\right] + \mathbb{E}^x\left[\int_0^\tau e^{-\lambda s}(\lambda - L) f(X_s)\, ds\right].$$

Boundary problem derivation

We can now form a boundary problem from the option pricing equation in (4.21),

$$V(T,x) = \mathbb{E}^x\big[e^{-r\tau_B} G_B(X_{\tau_B})\big],$$

by first returning to the general case in (4.5),

$$V(T,t,x) = \mathbb{E}^x\Big[e^{-r(\tau_B - t)} V(T, \widehat{X}_{\tau_B})\Big].$$

Consider Dynkin's formula for the function $f = V(T,t,x)$, stochastic process $\widehat{X}$, stopping time $\tau = \tau_B$, and $\lambda = r$. According to Lemma 10, the infinitesimal generator of $\widehat{X}$ is $\partial_t + L$, where L is the infinitesimal generator of X. Then,

$$V(T,t,x) = \mathbb{E}^x\Big[e^{-r\tau_B} V(T, \widehat{X}_{\tau_B})\Big] + \mathbb{E}^x\left[\int_0^{\tau_B} e^{-rs}(r - \partial_s - L)V(T, \widehat{X}_s)\, ds\right].$$

But, when V is evaluated in the region B, we have

$$V(T, \widehat{X}_t) = G_B(X_t),$$

where $G_B(X_t)$ is given in (4.22). Thus, since $\widehat{X}_{\tau_B} \in B$,

$$V(T,t,x) = \mathbb{E}^x\big[e^{-r\tau_B} G_B(X_{\tau_B})\big] + \mathbb{E}^x\left[\int_0^{\tau_B} e^{-rs}(r - \partial_s - L)V(T, \widehat{X}_s)\, ds\right]. \tag{4.23}$$

The first term on the right-hand side of (4.23) is precisely the expression in (4.21), so we need the second term to vanish. If $(t,x) \in B$, this occurs automatically since we have an integral from 0 to 0. We have previously seen that $V(T,t,x) = G_B(x)$ in this case. Alternatively, if $(t,x) \in C$, the required condition is equivalent to $(\partial_t + L - r)V(T,t,x) = 0$, as discussed in [8, Theorem 2.12]. These conclusions lead to the following boundary problem.

$$\begin{aligned} (\partial_t + L - r)V(T,t,x) &= 0, && (t,x) \in C, \\ V(T,t,x) &= G_B(x), && (t,x) \in B. \end{aligned}$$

Finally, we can decompose the second condition into the terminal and boundary conditions, giving the required boundary problem.

$$(\partial_t + L - r)V(T,t,x) = 0, \qquad x > h,\ t < T, \tag{4.24}$$

$$V(T,t,x) = 0, \qquad x \le h,\ t \le T, \tag{4.25}$$

$$V(T,T,x) = G_B(x), \qquad x > h. \tag{4.26}$$

Pricing method

The boundary problem in (4.24)-(4.26) can be solved using the *method of lines*, where we discretise time, such that $0 = t_0 < t_1 < \cdots < t_M = T$, for some $M \in \mathbb{N}$. While it is not required, suppose that this grid is uniform and set $q = (t_{s+1} - t_s)^{-1} = M/T$. Denote Carr's randomised approximation to $V(t,x)$ by $V^s(x)$ and replace the time derivative by a finite difference:

$$\partial_t V(t,x) \approx \frac{V^{s+1}(x) - V^s(x)}{t_{s+1} - t_s} = q\big(V^{s+1}(x) - V^s(x)\big).$$

Substituting the finite difference into the boundary problem and rearranging, we derive a new boundary problem. For $s = N-1, N-2, \ldots,$

$$\begin{aligned}(q - L)V^s(x) &= \frac{1}{q+r}V^{s+1}(x), \qquad && x > h,\\ V^s(x) &= 0, && x \le h.\end{aligned}$$

Carr's randomisation solves this boundary problem by assuming that the maturity time, T, is randomised according to a Gamma distribution with shape M and rate $q = M/T$, $\mathbf{g}(M,q)$, which converges to T in quadratic mean as M approaches infinity [33]. This provides the following probabilistic interpretation of this boundary problem:

$$V^s(x) \approx \mathbb{E}^x\big[e^{-r\mathbf{g}(M,q)}\, G(X_{\mathbf{g}(M,q)})\big].$$

This problem is solved in detail in [8], giving the same result as we found using the Laplace transform in Section 4.3.1. If

$$\begin{aligned}v_0(T,x) &= (K - e^x)^+ \mathbb{1}\{x > h\},\\ v_p(T,x) &= \frac{q}{q+r}\,\mathcal{F}^{-1}\,\phi^-_{q+r}(\xi)\,\mathcal{F}\,\mathbb{1}_{(h,\infty)}\,\mathcal{F}^{-1}\,\phi^+_{q+r}(\xi)\,\mathcal{F}\,v_{p-1}(T,x), \quad \text{for } p \in \mathbb{N}. \qquad (4.27)\end{aligned}$$

then $v_p(T,x)$ approaches $V(T,x)$ as M approaches infinity.

4.4 Simple Wiener-Hopf method

The Simple Wiener-Hopf method [33] is one of several algorithms that have been derived to numerically price barrier options using (4.20) (or, equivalently, (4.27)). These algorithms calculate the value of a barrier option where the underlying security is modelled by an exponential Lévy process, for certain subclasses of processes including jump processes and non-Gaussian distributions. The SWH method, and other similar algorithms, approximate the Wiener-Hopf factors which are not known for most of these Lévy processes. Below,

we summarise the SWH method and its assumptions and limitations, then implement it for comparison with our new algorithm.

The SWH method aims to split the Lévy process into the difference between two subordinators, which have relatively simple approximations for their Wiener-Hopf factors. This decomposition is possible if and only if the Lévy process is pure non-Gaussian and has jumps of finite variation [33], thus specifying the subclass of processes that may be used to drive the model. The process under consideration is separated into the sum of a non-decreasing process and a non-increasing process, which are formed from the positive and negative jumps, respectively, along with the drift term. If the drift is positive, it contributes to the non-decreasing process; otherwise, it contributes to the non-increasing process. These new processes are subordinators and admit trivial Wiener-Hopf factorisations, leading to simple option pricing formulae.

In what follows, consider a knock-out put barrier option with strike price K, barrier H, spot price S_0, and maturity T, and let $h \coloneqq \ln\left(\frac{H}{S_0}\right)$.

Trivial case 1

First, suppose that the log-price of the underlying security, $X^+ = \left\{X_t^+ \colon t \in \mathbb{R}_+\right\}$, is a subordinator that moves only by positive jumps of finite variation or positive linear drift. Thus, by (2.5), the characteristic exponent of X^+ is

$$\Psi_+(\xi) = -i\mu\xi + \int_{\mathbb{R}_+} (1 - e^{ix\xi})\, \Pi_+(dx),$$

where $\Pi_+(dx)$ is the Lévy measure of X^+. Clearly, X is almost surely non-decreasing, the supremum and infimum processes satisfy

$$\begin{aligned} \overline{X}_t &= X_t, \\ \underline{X}_t &= X_0 = 0 \quad \text{a.s.}. \end{aligned}$$

Finally, from Section 2.3.4, recall that the characteristic function of $X_{\mathbf{e}_{q+r}}$ is $\frac{q+r}{q+r+\Psi_+(\xi)}$. Also, notice that the characteristic function of a zero constant is 1. Thus, the characteristic functions of $\overline{X}_{\mathbf{e}_{q+r}}$ and $\underline{X}_{\mathbf{e}_{q+r}}$ are

$$\begin{aligned} \phi^+_{q+r}(\xi) &= \frac{q+r}{q+r+\Psi_+(\xi)}, \\ \phi^-_{q+r}(\xi) &= 1, \end{aligned}$$

respectively. Consider the application of the Laplace transform method from Section 4.3.1 in this situation. Then, (4.18) can be rewritten as

$$\begin{aligned} v_p(T,x) &= \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}\big(T,\, X_{\mathbf{e}_{q+r}}\big)\,\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big] \\ &= \frac{q\,\mathbb{1}\{x>h\}}{q+r}\,\mathbb{E}^x\big[v_{p-1}\big(T,\, \overline{X}_{\mathbf{e}_{q+r}}\big)\big] \\ &= \frac{q\,\mathbb{1}\{x>h\}}{q+r}\,\mathcal{E}^+_{q+r}\,v_{p-1}(T,x), \end{aligned}$$

where $\mathcal{E}^+_{q+r}$ is the expected present value (EPV) operator for $\overline{X}_{\mathbf{e}_{q+r}}$, as defined in (4.12). Applying Lemma 9 to the EPV operator, we have

$$v_p(T,x) = \frac{q\,\mathbb{1}\{x>h\}}{q+r}\,\frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}\,\frac{q+r}{q+r+\Psi_+(\xi)}\,\widehat{v}_{p-1}(q,\xi)\,d\xi,$$

where $\widehat{v}_{p-1}(q,x)$ is the Fourier transform of $v_{p-1}(T,x)$ with respect to T. We now have an iterative formula that converges to the option price, as was derived using the Laplace transform and Carr's randomisation techniques previously. However, notice that in this trivial example where the Lévy process is a subordinator, we do not have to calculate the Wiener-Hopf factors. Instead, we simply calculate $\Psi_+(\xi)$, which is much more commonly available.

Trivial case 2

Alternatively, suppose that the log-price of the underlying security, $X^- = \{X^-_t : t \in \mathbb{R}_+\}$, has only negative jumps of finite variation and negative linear drift. This means that X is non-increasing and so, the negative process, $-X^-$, is a subordinator. Similar to previously, the supremum and infimum processes of X are defined by

$$\begin{aligned} \overline{X}_t &= X_0 = 0 \quad \text{a.s.}, \\ \underline{X}_t &= X_t, \end{aligned}$$

and the characteristic functions of $\overline{X}_{\mathbf{e}_{q+r}}$ and $\underline{X}_{\mathbf{e}_{q+r}}$ are

$$\begin{aligned} \phi^+_{q+r}(\xi) &= 1, \\ \phi^-_{q+r}(\xi) &= \frac{q+r}{q+r+\Psi_-(\xi)}, \end{aligned}$$

respectively, where Ψ_- is the characteristic exponent of X^-, with the same form as Ψ_+ when $\Pi_+(dx)$ is replaced by $\Pi_-(dx)$. Working with (4.18) again, in this case we have

$$\begin{aligned} v_p(T,x) &= \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}(T,\, X_{\mathbf{e}_{q+r}})\,\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big] \\ &= \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}(T,\, \underline{X}_{\mathbf{e}_{q+r}})\,\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big] \\ &= \frac{q}{q+r}\,\mathcal{E}^-_{q+r}\,\mathbb{1}_{(h,\infty)}\,v_{p-1}(T,\,x), \end{aligned}$$

where $\mathcal{E}^-_{q+r}$ is the EPV operator for $\underline{X}_{\mathbf{e}_{q+r}}$ and $\mathbb{1}_{(h,\infty)}$ is an operator form of the indicator function where $\mathbb{1}_{(h,\infty)}u(x) = u(\mathbb{1}\{x > h\})$. Using Lemma 9 as before,

$$v_p(T,x) = \frac{q}{q+r}\,\frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}\,\frac{q+r}{q+r+\Psi_-(\xi)}\,\widehat{v^*_{p-1}}(q,\xi)\,d\xi,$$

where $v^*_{p-1} = \mathbb{1}_{(h,\infty)}\,v_{p-1}(T,x)$. Again, if the Lévy processes is almost surely non-increasing, we do not need to calculate the Wiener-Hopf factors and can use a much more accessible formula.

General case

More generally, suppose that the log-price of the underlying security, $X = \{X_t\colon t \in \mathbb{R}_+\}$, is purely non-Gaussian and has finite jump activity. Then X can be decomposed into the difference of two subordinators, $X^+ = \{X^+_t\colon t \in \mathbb{R}_+\}$ and $-X^- = \{-X^-_t\colon t \in \mathbb{R}_+\}$, with

$$X_t = X^+_t - (-X^-_t) \quad \forall\, t \in \mathbb{R}_+.$$

Let Ψ_+ and Ψ_- be the characteristic exponents of X^+ and X^-, respectively. Observe that, for all $t \in \mathbb{R}_+$,

$$\overline{X}^+_t = X^+_t, \tag{4.28}$$
$$\underline{X}^+_t = X^-_t. \tag{4.29}$$

Let $X^{+,1} = \{X^{+,1}_t\colon t \in \mathbb{R}_+\}$ and $X^{+,2} = \{X^{+,2}_t\colon t \in \mathbb{R}_+\}$ be independent Lévy processes that both have the same characteristic exponent as X^+, so $X^{+,1}_t \overset{d}{=} X^{+,2}_t \overset{d}{=} X^+_t$. Then, by (2.6) and the independent and stationary increment properties of X, for all $t \in \mathbb{R}_+$,

$$X_t \overset{d}{=} \overline{X}_t - \underline{X}_t \overset{d}{=} X^{+,1}_{t/2} + X^-_t + X^{+,2}_{t/2}.$$

Given $t \in \mathbb{R}_+$ and $X_t = x$, define another process, $Y = \{Y_s\colon 0 \le s \le 2t\}$ such that

i. Y starts at x: $Y_0 = x$,

ii. Y moves upward according to X^+_t for $s \in [0,\, t/2)$: $Y_s = x + X^{+,1}_s$,

iii. Y moves downward according to X^-_t for $s \in [t/2,\, 3t/2)$: $Y_s = Y_{t/2} + X^-_{s-t/2}$,

iv. Y moves upward according to X^+_t for $s \in [3t/2,\, 2t]$: $Y_s = Y_{3t/2} + X^{+,2}_{s-3t/2}$.

Then, according to [33], for short time periods $[0, t]$,

$$\overline{X}_t \approx \overline{Y}_{2t} \quad \text{and} \quad \underline{X}_t \approx \underline{Y}_{2t}.$$

The dynamics of the supremum and infimum processes of X are exactly what we require to solve the option pricing formulae that we saw in Sections 4.3.1 and 4.3.2. These extrema are given by the following expressions.

$$\overline{Y}_{2t} = \max\Big\{x,\, x + X^{+,1}_{t/2},\, x + X^{+,1}_{t/2} + X^{-}_{t} + X^{+,2}_{t/2}\Big\},$$
$$\underline{Y}_{2t} = \min\Big\{x,\, x + X^{+,1}_{t/2} + X^{-}_{t}\Big\}.$$

Let $M \in \mathbb{N}$ be sufficiently large and define $q = M/T$. Recall in (4.18), we defined $v_p(T,x)$ that converges to $V(T,x)$, where the processes in these equations were evaluated at the random time $\mathbf{e}_{q+r}$. As M approaches infinity, so does q and, hence, $\mathbf{e}_{q+r}$ approaches 0. This means that we can approximate $X_{\mathbf{e}_{q+r}}$ with $Y_{2\mathbf{e}_{q+r}}$ in these calculations.

We will now complete the derivation in a similar way to the Laplace transform method in Section 4.3.1, but with this new process which trivialises the Wiener-Hopf factors. Observe the following relationships.

$$Y_{2\mathbf{e}_{q+r}} = x + X^{+,1}_{\mathbf{e}_{q+r}/2} + X^{-}_{\mathbf{e}_{q+r}} + X^{+,2}_{\mathbf{e}_{q+r}/2}, \tag{4.30}$$
$$\mathbb{1}\{\underline{Y}_{2\mathbf{e}_{q+r}} > h\} = \mathbb{1}\{x > h\}\,\mathbb{1}\{x + X^{+,1}_{\mathbf{e}_{q+r}/2} + X^{-}_{\mathbf{e}_{q+r}}\}. \tag{4.31}$$

The EPV operators are defined similarly to previously, but since Y_t involves $X_{t/2}$, not X_t, the positive EPV operator involves this scaled time. Scaling an exponential variable is equivalent to scaling the rate parameter, so $\mathbf{e}_{q+r}/2 \overset{d}{=} \mathbf{e}_{2(q+r)}$. Thus, define

$$\mathcal{E}^{+}u(x) \coloneqq \mathbb{E}\Big[u\big(x + \overline{X}^{+}_{\mathbf{e}_{2(q+r)}}\big)\Big] = \mathbb{E}\Big[u\big(x + X^{+}_{\mathbf{e}_{2(q+r)}}\big)\Big], \tag{4.32}$$
$$\mathcal{E}^{-}u(x) \coloneqq \mathbb{E}\Big[u\big(x + \underline{X}^{-}_{\mathbf{e}_{q+r}}\big)\Big] = \mathbb{E}\Big[u\big(x + X^{-}_{\mathbf{e}_{q+r}}\big)\Big], \tag{4.33}$$

using (4.28) and (4.29). Importantly, since these processes are subordinators, the trivial cases that we investigated earlier show that their characteristic functions (or, equivalently, their Wiener-Hopf factors) are

$$\phi^{+}_{2(q+r)}(\xi) \coloneqq \mathbb{E}\Big[e^{i\xi X^{+}_{\mathbf{e}_{2(q+r)}}}\Big] = \frac{2(q+r)}{2(q+r) + \Psi_{+}(\xi)}, \tag{4.34}$$
$$\phi^{-}_{q+r}(\xi) \coloneqq \mathbb{E}\Big[e^{i\xi X^{-}_{\mathbf{e}_{q+r}}}\Big] = \frac{q+r}{q+r+\Psi_{-}(\xi)}. \tag{4.35}$$

As in (4.12) and (4.13), we can write these EPV operators using Fourier transforms which may be efficiently implemented numerically with the fast Fourier transform (Appendix A.2):

$$\mathcal{E}^{+}_{2(q+r)}u(x) = \mathcal{F}^{-1}\,\frac{2(q+r)}{2(q+r)+\Psi_{+}(\xi)}\,\mathcal{F}\,u(x), \tag{4.36}$$
$$\mathcal{E}^{-}_{q+r}u(x) = \mathcal{F}^{-1}\,\frac{q+r}{q+r+\Psi_{-}(\xi)}\,\mathcal{F}\,u(x). \tag{4.37}$$

Finally, let $Z = X^{+,1}_{\mathbf{e}_{q+r}/2} + X^{-}_{\mathbf{e}_{q+r}}$ and $W = X^{+,1}_{\mathbf{e}_{q+r}/2}$. Then, following from (4.18), for $q = M/T$ with M sufficiently large,

$$\begin{aligned}
v_p(T,x) &= \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}\big(T,\, X_{\mathbf{e}_{q+r}}\big)\,\mathbb{1}\{\underline{X}_{\mathbf{e}_{q+r}} > h\}\Big] \\
&\approx \frac{q}{q+r}\,\mathbb{E}^x\Big[v_{p-1}(T,\, Y_{\mathbf{e}_{2(q+r)}})\,\mathbb{1}\{\underline{Y}_{\mathbf{e}_{2(q+r)}} > h\}\Big] \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathbb{E}^x\Big[v_{p-1}(T,\, Z + X^{+,2}_{\mathbf{e}_{2(q+r)}})\,\mathbb{1}\{Z>h\}\Big] && \text{(by (4.30))} \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathbb{E}^x\Big[\mathbb{1}\{Z>h\}\,v_{p-1}(T,\, W + X^{-}_{\mathbf{e}_{q+r}} + X^{+,2}_{\mathbf{e}_{2(q+r)}})\Big] \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathbb{E}^x\Big[\mathbb{1}\{Z>h\}\,\mathcal{E}^+_{2(q+r)}\,v_{p-1}(T,\, W + X^{-}_{\mathbf{e}_{q+r}})\Big] && \text{(by (4.36))} \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathbb{E}^x\Big[\mathbb{1}\{W + X^{-}_{\mathbf{e}_{q+r}} > h\}\,\mathcal{E}^+_{2(q+r)}\,v_{p-1}(T,\, W + X^{-}_{\mathbf{e}_{q+r}})\Big] \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathbb{E}^x\Big[\mathcal{E}^-_{q+r}\,\mathbb{1}\{W>h\}\,\mathcal{E}^+_{2(q+r)}\,v_{p-1}(T,\, W)\Big] && \text{(by (4.37))} \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathbb{E}\Big[\mathcal{E}^-_{q+r}\,\mathbb{1}\{x + X^{+,1}_{\mathbf{e}_{2(q+r)}} > h\}\,\mathcal{E}^+_{2(q+r)}\,v_{p-1}(T,\, x + X^{+,1}_{\mathbf{e}_{2(q+r)}})\Big] \\
&= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathcal{E}^+_{2(q+r)}\,\mathcal{E}^-_{q+r}\,\mathbb{1}_{(h,\infty)}\,\mathcal{E}^+_{2(q+r)}\,v_{p-1}(T,\, x). && \text{(by (4.36))}
\end{aligned}$$

The SWH method defines approximations, in (4.34) and (4.35), to the Wiener-Hopf factors of a non-Gaussian Lévy process with finite jump activity. This allows us to define the EPV operators in (4.36) and (4.37) and calculate $v_p(T,x)$ iteratively, with and

$$\begin{aligned}
v_0(T,x) &= (K - e^x)^+\mathbb{1}\{x>h\}, \\
v_p(T,x) &= \frac{\mathbb{1}\{x>h\}(x)}{1+r/q}\,\mathcal{E}^+_{2(q+r)}\,\mathcal{E}^-_{q+r}\,\mathbb{1}_{(h,\infty)}\,\mathcal{E}^+_{2(q+r)}\,v_{p-1}(T,\, x), \quad \text{for } p \in \mathbb{N}.
\end{aligned} \tag{4.38}$$

The required option value, $V(T,x)$, is approximated by $v_M(T,x)$ for large M, giving the option pricing model.

4.4.1 SWH method implementation

We now implement the SWH method in Python to demonstrate the algorithm and its accuracy. The underlying Lévy process is initially assumed to be a KoBoL process, described in Section 4.1.1, which is the same process used in the paper that originally introduced the SWH method [33]. We also use the same parameters such that the paper can be used to verify our outputs.

Consider a financial security with initial spot price S_0, ranging from 91 to 121, in a market with risk-free interest rate $r = 0.07231$. Suppose that the log-price of the security follows a KoBoL process with parameters $\nu_+ = \nu_- = 0.5$, $c_+ = c_- = 1$, $\lambda_+ = 9$, and $\lambda_- = -8$. Consider a knock-out barrier put option for this security with strike price $K = 100$, barrier $H = 90$, and maturity $T = 0.5$ years. We implement the SWH method as follows, starting by inputting the constants.

```
import numpy as np
from scipy.fft import fft, ifft
from scipy.special import gamma
pi = np.pi

# KoBoL process constants
nu_plus = nu_minus = 0.5
c_plus = c_minus = 1
lambda_plus = 9
lambda_minus = -8

# Option constants
K = 100
S0 = 91
H = 90
r = 0.07231
T = 0.5
```

Then, the drift, μ, is calculated according to (4.2), giving a value of $\mu \approx 0.0358$.

```
def lower_func(x, nu, lambda_plus):
    if (-1 <= x) & (x <= 1):
        indicator = 1
    else:
        indicator = 0
    return (1 - np.exp(x) + x*indicator) * (abs(x)**(-nu - 1) *
            np.exp(lambda_plus * x))

def upper_func(x, nu, lambda_minus):
    if (-1 <= x) & (x <= 1):
        indicator = 1
    else:
        indicator = 0
    return (1 - np.exp(x) + x*indicator) * (abs(x)**(-nu - 1) *
            np.exp(lambda_minus * x))

I_lower = quad(lower_func, -np.inf, 0, args=(nu, lambda_plus))
I_upper = quad(upper_func, 0, np.inf, args=(nu, lambda_minus))
mu = r + sigma**2/2 + c*(I_lower[0] + I_upper[0])
```

Next, we define additional constants: $h = \ln(H/S_0)$, the barrier in the log-domain; $N_{pw} = 800$, the number of Post-Widder samples; and $q = N_{pw}/T$, the observation rate.

```
h = np.log(H/S0)      # Log-barrier price
N_pw = 800            # Number of Post-Widder samples
q = N_pw / T          # Observation rate
```

There are two domains on which the variables are defined: the space domain (the log-price domain of the security) and the frequency domain (the target of the Fourier transform). The space domain step size, *xstep*, is a parameter of the algorithm that is negatively correlated with both its accuracy and evaluation speed (*i.e.* smaller step size leads to higher accuracy and required time). The number of steps, M, is the same in both domains and must be at least large enough so that the space domain covers the value $x = 1$. Then, the frequency domain step size is given by $2\pi/(xstep * M)$. Suggested values of $xstep = 0.001$ and $M = 2^{10}$ are given in [32] and implemented as follows.

```
M = 2**10                     # Number of steps
xstep = 0.001                 # Space domain step size
x = np.zeros(M+1)             # Space domain grid points
for k in range(0, M+1):
    x[k] = -M*xstep/2 + k*xstep
xi = np.zeros(M+1)            # Frequency domain grid points
for l in range(-int(M / 2), int(M / 2)):
    xi[l] = 2 * pi * l / (xstep * M)
```

The next step in the SWH method is to define the characteristic exponents for the two subordinators, X^+ and $-X^-$, where X is the underlying KoBoL process and

$$X = X^+ - (-X^-).$$

Since a KoBoL process is a pure jump process with drift, we can decompose it by splitting the process into its positive and negative jumps, with the drift added to the corresponding process depending on its sign. By Lemma 7, if $\mu > 0$, then the characteristic exponents of X^+ and X^- are

$$\Psi^+(\xi) = -i\mu\xi + c_+\Gamma(-\nu_+)[\lambda_+^{\nu_+} - (\lambda_+ + i\xi)^{\nu_+}],$$
$$\Psi^-(\xi) = c_-\Gamma(-\nu_-)[(-\lambda_-)^{\nu_-} - (-\lambda_- - i\xi)^{\nu_-}],$$

respectively. Otherwise,

$$\Psi^+(\xi) = c_+\Gamma(-\nu_+)[\lambda_+^{\nu_+} - (\lambda_+ + i\xi)^{\nu_+}],$$
$$\Psi^-(\xi) = -i\mu\xi + c_-\Gamma(-\nu_-)[(-\lambda_-)^{\nu_-} - (-\lambda_- - i\xi)^{\nu_-}].$$

These characteristic exponents are defined in our code below by the functions **psi_plus** and **psi_minus**. These functions are used to calculate the approximate Wiener-Hopf factors, **phi_plus** and **phi_minus**, according to (4.34) and (4.35).

```
def Psi_plus(xi):
    return -1j*xi*[mu if mu > 0 else 0][0] + c_plus * gamma(-nu_plus) *
    (lambda_plus ** nu_plus - (lambda_plus + 1j * xi) ** nu_plus)

def Psi_minus(xi):
    return -1j*xi*[mu if mu < 0 else 0][0] + c_minus * gamma(-nu_minus) *
    ((-lambda_minus) ** nu_minus - (-lambda_minus - 1j * xi) ** nu_minus)

phi_plus = np.zeros(len(xi), dtype=complex)
phi_minus = np.zeros(len(xi), dtype=complex)
for ixi in range(len(xi)):
    phi_plus[ixi] = (2 * (q + r)) / (2 * (q + r) + Psi_plus(xi[ixi]))
    phi_minus[ixi] = (q + r) / (q + r + Psi_minus(xi[ixi]))
```

Finally, the payoff function is defined in **calc_G** so that the initial term of the iteration formula, $v_0 = (K - S_0 e^x)^+$, can be calculated. Then, after defining the required indicator function, $\mathbb{1}\{x > h\}$, as an array for each x in the domain, we iteratively evaluate $v_p(T, x)$ according to (4.38) and return the calculated option value for the initial security price $S = S_0 e^0 = S_0$, which occurs when $x = 1$.

```
def calc_G(K, S0, x):
    return max(K - S0*np.exp(x), 0)

v = np.zeros((N_pw, len(x)))
for ix in range(len(x)):
    v[0, ix] = calc_G(K, S0, x[ix])

indicatorvec = np.zeros(len(x))
for i in range(len(indicatorvec)):
    if x[i] > h:
        indicatorvec[i] = 1

for i in range(1, N_pw):
    v[i] = (1 / (1 + r / q)) *
        np.multiply(indicatorvec,
        ifft(np.multiply(phivals_plus, fft(
        ifft(np.multiply(phivals_minus, fft(
        np.multiply(indicatorvec,
        ifft(np.multiply(phivals_plus, fft(v[i - 1, :])))))))))))

# Output option value estimate from final iteration
V = v[-1, :]
V_out = V[min(enumerate(np.exp(x)), key=lambda x: abs(x[1] - 1))[0]]
```

Table 4.1: SWH results for a knock-out barrier put option with $K = 100$, $H = 90$, $r = 0.07231$, and $T = 0.5$, driven by a KoBoL process with $\mu = 0$, $\nu_+ = \nu_- = 0.5$, $c_+ = c_- = 1$, $\lambda_+ = 9$, and $\lambda_- = -8$.

	Monte Carlo	**SWH paper**	**Our implementation**
$\mathbf{N_{pw}}$	10^5	800	800
$\mathbf{S_0 = 91}$	0.23650	0.23652	0.23808
$\mathbf{S_0 = 101}$	0.56997	0.56571	0.57226
$\mathbf{S_0 = 111}$	0.38399	0.38388	0.38767
$\mathbf{S_0 = 121}$	0.20949	0.20974	0.21023
$\mathbf{S_0 = 131}$	0.10836	0.10805	0.11211

Table 4.1 shows the output of our implementation of the SWH method, along with the results given in [33], which included both the SWH method and a Monte Carlo simulation. Despite some minor differences in the results, Table 4.1 suggests that our implementation is reasonably accurate.

4.5 Wiener-Hopf Monte Carlo simulation of Lévy processes

The accuracy of numerical algorithms should be evaluated by comparison with expected outcomes. This is often achieved using Monte Carlo simulation, where many realisations of the underlying stochastic process are simulated and averaged. In some cases, such as with Brownian motion, this is a relatively simple endeavour. However, simulating general Lévy processes requires alternative methods, which have been discussed in the literature (see, *e.g.*, [31, 37, 56]). Here, we summarise and implement a Wiener-Hopf Monte Carlo (WHMC) method by Kuznetsov [37], which we will later use as a comparison with our algorithm. The idea for the WHMC algorithm follows from the Carr's randomisation techniques that we used previously [11]. A Lévy process, $X = \{X_t \colon t \in \mathbb{R}_+\}$, is simulated on a grid that is randomised according to exponential realisations. Then, the Wiener-Hopf factorisation is used to determine the distributions of X and $\overline{X}$ and, hence, sample the path of the process.

Suppose that the Lévy process, X, is defined on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we want to simulate it on the time interval $[0, t]$. Define a sequence of i.i.d. exponentially distributed random variables with unit mean, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_M$, and product law $\mathbf{P}$ orthogonal to $(\Omega, \mathcal{F}, \mathbb{P})$. These variables form the grid on which we simulate X, so we need to scale them such that the expected simulation length is t. The variables need to be scaled by t/n since, by the strong law of large numbers,

$$\lim_{n\to\infty} \sum_{j=1}^{M} \frac{t}{n} \mathbf{e}_j = t \quad \mathbf{P}\text{-a.s..} \tag{4.39}$$

This sum of exponential random variables is, by definition, equal in law to a Gamma random variable with parameters M and M/t, $\mathbf{g}(n,\, n/t)$. Kuznetsov [37] argues that, for sufficiently large n, the joint probability of X_t and $\overline{X}_t$ can be approximated using the joint probability at the random time $\mathbf{g}(n,\, n/t)$:

$$\mathbb{P}(X_t \in dx,\, \overline{X}_t \in dy) \approx (\mathbf{P} \times \mathbb{P})(X_{\mathbf{g}(n,\, n/t)} \in dx,\, \overline{X}_{\mathbf{g}(n,n/t)} \in dy).$$

The distribution on the right-hand side of this equation is concerned with the path of the Lévy process over exponential time periods, meaning that we can apply the Wiener-Hopf factorisation to derive Theorem 10.

Theorem 10. *Let $n \in \mathbb{N}$, $\lambda > 0$, and define $\mathbf{g}(n,\, \lambda) \coloneqq \sum_{j=1}^{n} \mathbf{e}_i/\lambda$. Let $S_0^{(\lambda)} = I_0^{(\lambda)} \coloneqq 0$, and $\{S_j^{(\lambda)} \colon j \geq 1\}$ and $\{I_j^{(\lambda)} \colon j \geq 1\}$ be i.i.d. sequences of random variables with distribution equal to $\overline{X}_{\mathbf{e}_1/\lambda}$ and $\underline{X}_{\mathbf{e}_1/\lambda}$, respectively. Define $A_n^{(\lambda)}$ and $B_n^{(\lambda)}$ iteratively such that $A_0^{(\lambda)} = B_0^{(\lambda)} \coloneqq 0$ and, for $j = 1, 2, \ldots, n$,*

$$\begin{aligned} A_j^{(\lambda)} &\coloneqq A_{j-1}^{(\lambda)} + S_j^{(\lambda)} + I_j^{(\lambda)}, \\ B_j^{(\lambda)} &\coloneqq \max\big(B_{j-1}^{(\lambda)},\, A_{j-1}^{(\lambda)} + S_j^{(\lambda)}\big). \end{aligned}$$

Then,

$$(X_{\mathbf{g}(n,\, \lambda)},\, \overline{X}_{\mathbf{g}(n,\, \lambda)}) \overset{d}{=} (A_n^{(\lambda)},\, B_n^{(\lambda)}),$$

Proof. See [37, Theorem 1]. □

Thus, provided that we can sample from $\overline{X}_{\mathbf{e}_1/\lambda}$ and $\underline{X}_{\mathbf{e}_1/\lambda}$, we now have an iterative method of calculating the joint distribution of X and $\overline{X}$ at the gamma random time $\mathbf{g}(n, \lambda)$. Letting $\lambda = n/t$, this variable approaches t as n approaches infinity. So,

$$\lim_{n\to\infty} \left(A_n^{(\frac{n}{t})},\, B_n^{(\frac{n}{t})}\right) = (X_t,\, \overline{X}_t).$$

This equality allows us to estimate functions of $(X_t,\, \overline{X}_t)$ by using standard Monte Carlo methods with $(A_n^{(\frac{n}{t})},\, B_n^{(\frac{n}{t})})$ for sufficiently large n. Given a function F, simulate i.i.d. copies of $(A_n^{(\frac{n}{t})},\, B_n^{(\frac{n}{t})})$, denoted $(A_{n,\,k}^{(\frac{n}{t})},\, B_{n,\,k}^{(\frac{n}{t})})$, for $k = 1, 2, \ldots, m$ and then, for sufficiently large m,

$$\mathbb{E}\big[F(X_t,\, \overline{X}_t)\big] \approx \frac{1}{m} \sum_{k=1}^{m} F\Big(\Big(A_{n,\,k}^{(\frac{n}{t})},\, B_{n,\,k}^{(\frac{n}{t})}\Big)\Big).$$

For our purposes, the function, F, is the discounted payout function of the option under consideration. This Wiener-Hopf Monte Carlo method is several orders of magnitude slower than the previous option pricing procedures, but it allows us to calculate a target value for our algorithms. The choice of methods to sample $A_n^{(\lambda)}$ and $B_n^{(\lambda)}$ depends on the particular Lévy process that is being used, and will be discussed later.

Remark 1. *The techniques that we used to build our Lévy-driven models, both the Laplace transform method and Carr's randomisation, relied on an observation rate q. This parameter defined the rate of exponential realisations at which our algorithm was observed. Similarly, the WHMC method was also defined on an exponential grid, with parameter λ. In both cases, the exponential realisations can alternatively be interpreted as the arrival times of a Poisson process. Thus, the algorithms in this thesis are observed at Poisson arrival times, as alluded to in the title.*

Remark 2. *As the rate parameter of a Poisson process tends to infinity, the time between arrivals approaches 0. Accordingly, as the parameter λ increases, our algorithm approaches the continuously-observed result, and thus, accuracy increases. This will however also lead to increased computational time.*

Chapter 5

Meromorphic Lévy processes

We have seen that Lévy processes may be used to model option prices by utilising the Wiener-Hopf factorisation. While there are various approximation techniques, such as the Simple Wiener-Hopf (SWH) method, it would be desirable to have models driven by processes that have explicit expressions for their Wiener-Hopf factors. One such example is the class of *meromorphic Lévy processes*, which was introduced in [34, 35]. In this section, we discuss these processes and their characteristics, with the aim of eventually implementing them into our option pricing formulae.

5.1 Meromorphic Lévy processes

Meromorphic Lévy processes are a class of Lévy processes that were derived largely because they admit explicit expressions for their Wiener-Hopf factors. The conception of this class began by considering the case of a Lévy process that is the difference of two independent compound Poisson processes. In this case, the Wiener-Hopf factors may be tractable because the characteristic exponent is a rational function, the ratio of two polynomials of finite degree, which can easily be decomposed into Wiener-Hopf factors, as discussed in [38, Section 6.5.4]. Various authors have attempted to define classes of Lévy processes with tractable Wiener-Hopf factors by choosing different jump distributions for the compound Poisson processes. Meromorphic Lévy processes are defined accordingly, with jump distribution given by an infinite mixture of exponentials. The following derivation of the meromorphic class of Lévy processes was due to Kuznetsov in [34, 35, 37]. We first define *completely monotone functions* and then *discrete completely monotone functions*, which are shown to be intimately tied to an infinite mixture of exponentials.

Definition 41 (Completely monotone function)**.** *A function* $f\colon (0,\infty) \to \mathbb{R}$ *is* ***completely monotone*** *if* $f \in C^\infty$ *and, for all* $n \in \mathbb{Z}_+$ *and* $x > 0$*,*

$$(-1)^n f^{(n)}(x) \geq 0.$$

An important related class is *monotonic functions*, where only the first derivative needs to satisfy the given condition; the set of completely monotone functions is a subset of the monotonic functions. Some examples of completely monotone functions are x^α for $0 \leq \alpha \leq 1$, $\log(1+x)$, and $1 - e^{-tx}$ for $t > 0$, as can be verified directly from the definition. The Bernstein theorem is often used to provide another interpretation of completely monotone functions, and leads to the subclass of discrete completely monotone functions.

Theorem 11 (Bernstein theorem)**.** *A function,* $f\colon (0,\infty) \to \mathbb{R}_+$*, is called completely monotone if and only if it is equal to the Laplace transform of a unique positive measure,* η*, on* $\mathbb{R}_+$*:*

$$f(x) = \int_{\mathbb{R}_+} e^{-qx}\eta(dq), \quad \forall\, x > 0. \tag{5.1}$$

The properties of the measure $\eta(dq)$ allow us to categorise functions into classes, such as discrete completely monotone functions.

Definition 42 (Discrete completely monotone)**.** *A completely monotone function* $f\colon (0,\infty) \to \mathbb{R}$ *is called* ***discrete completely monotone*** *if the measure* $\eta(dq)$ *in the Bernstein theorem is discrete, its support is infinite and does not have finite limit points.*

The conditions for a discrete completely monotone function imply that the measure, $\eta(dq)$, is an infinite mixture of atoms. Letting these atoms have size $a_n \geq 0$ and positions $b_n \geq 0$,

$$\eta(dq) = \sum_{n\in\mathbb{N}} a_n \delta_{b_n}(dq), \tag{5.2}$$

where δ_{b_n} is the Dirac measure at b_n. Since the support of $\eta(dq)$ is infinite and without finite limit points, we must have $b_n \to \infty$ as n approaches infinity. Without loss of generality, we could rearrange b_n such that the set $\{b_n\}_{n\in\mathbb{N}}$ is strictly increasing. Substituting (5.2) into Equation (5.1) from the Bernstein theorem shows that a discrete completely monotone function can be represented as an infinite series of exponential functions

$$f(x) = \int_{\mathbb{R}_+} e^{-qx}\,\eta(dq) = \sum_{n\in\mathbb{N}} a_n e^{-b_n x}, \quad x > 0. \tag{5.3}$$

As initially stated, we want to define the class of meromorphic Lévy processes according to a mixture of exponentials, which corresponds to having the *tails* of our Lévy measure, $\overline{\Pi}^+(x) = \Pi((x,\infty))$ and $\overline{\Pi}^-(x) = \Pi((-\infty,-x))$, be discrete completely monotone functions.

Definition 43 (Meromorphic Lévy process)**.** *Let X be a Lévy process with Lévy triple (μ, σ, Π). Then, X is a* ***meromorphic Lévy process*** *if $\overline{\Pi}^+$ and $\overline{\Pi}^-$ are discrete completely monotone functions.*

From (5.3)*, X is meromorphic if and only if the Lévy measure $\Pi(dx)$ has density with respect to the Lebesgue measure,*

$$\nu(x) = \mathbb{1}\{x > 0\} \sum_{n\in\mathbb{N}} a_n \rho_n e^{-\rho_n x} + \mathbb{1}\{x < 0\} \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n e^{\widehat{\rho}_n x}, \tag{5.4}$$

where a_n, $\widehat{a}_n$ are non-negative, ρ_n, $\widehat{\rho}_n$ are positive, $\{\rho_n\}_{n\in\mathbb{N}}$ and $\{\widehat{\rho}_n\}_{n\in\mathbb{N}}$ are strictly increasing with $\lim_{n\to\infty} \rho_n = \lim_{n\to\infty} \widehat{\rho}_n = \infty$, and these variables satisfy the summability condition

$$\sum_{n\in\mathbb{N}} a_n \rho_n^{-2} + \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n^{-2} < \infty. \tag{5.5}$$

The summability condition (5.5) is required to ensure that the series in (5.4) converges (Lemma 11) and the integrability condition for the Lévy measure, $\int_{\mathbb{R}} (1 \wedge x^2)\, \Pi(dx) < \infty$, is satisfied (Lemma 13). The summability condition is also equivalent to another integrability condition given in Lemma 12. Note that the integrands and summands throughout this section are all positive, allowing us to separate convergent integrals and use comparison tests in obvious ways.

Lemma 11. *The summability condition for a meromorphic Lévy process,* (5.5)*, implies that the series that defines the Lévy measure density* (5.4) *converges uniformly for $x \in \mathbb{R}\backslash\{0\}$.*

Proof. Suppose that

$$\sum_{n\in\mathbb{N}} a_n \rho_n^{-2} + \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n^{-2} < \infty.$$

Then,

$$\sum_{n\in\mathbb{N}} a_n \rho_n^{-2} < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n^{-2} < \infty.$$

Let $x \in \mathbb{R}_+\backslash\{0\}$. Since $\{\rho_n\}_{n\in\mathbb{N}}$ is positive and strictly increasing, choose $N \in \mathbb{N}$ such that $\rho_N > x$. Then, for all $n \geq N$, $e^{-\rho_n x} < e^{-x^2} < x^{-2} < \rho_n^{-2}$. Thus,

$$\sum_{n=N}^{\infty} a_n \rho_n e^{-\rho_n x} < \sum_{n=N}^{\infty} a_n \rho_n^{-2} < \sum_{n\in\mathbb{N}} a_n \rho_n^{-2} < \infty.$$

By the Weierstrass M-test, $\sum_{n=N}^{\infty} a_n \rho_n e^{-\rho_n x}$ converges uniformly. Since the summands are all finite, this implies that

$$\nu(x) = \sum_{n \in \mathbb{N}} a_n \rho_n e^{-\rho_n x} < \infty.$$

Analogous reasoning applies with $\widehat{\rho}_n$ if $x < 0$. Thus, the Lévy measure density converges uniformly for $x \in \mathbb{R} \backslash \{0\}$. □

Lemma 12. *The summability condition for a meromorphic Lévy process,* (5.5)*, is satisfied if and only if* $\int_{\mathbb{R}} x^2 \, \Pi(dx) < \infty$.

Proof. See [36, Proposition 1]. □

Lemma 13. *The summability condition,* (5.5)*, implies that the integrability condition for the Lévy measure,*

$$\int_{\mathbb{R}} (1 \wedge x^2) \, \Pi(dx) < \infty,$$

is satisfied.

Proof. Suppose that the summability condition is satisfied. Then, by Lemma 12,

$$\int_{\mathbb{R}} x^2 \, \Pi(dx) < \infty.$$

Since $(1 \wedge x^2) \leq x^2$ for all $x \in \mathbb{R}$,

$$\int_{\mathbb{R}} (1 \wedge x^2) \, \Pi(dx) \leq \int_{\mathbb{R}} x^2 \, \Pi(dx) < \infty.$$

Thus,

$$\int_{\mathbb{R}} (1 \wedge x^2) \, \Pi(dx) < \infty.$$

□

Now that we have defined meromorphic Lévy processes, we can discuss various properties of the class that may be beneficial in our option pricing model. Ideas for the derivations in this section are based on [36, Theorem 1]. We include modified proofs since we have slightly different notations and definitions.

5.1.1 Characteristic exponent

Let $X = \{X_t : t \in \mathbb{R}_+\}$ be a meromorphic Lévy process with Lévy triple (μ, σ, Π), meaning that $\Pi(dx)$ has density given by (5.4). Recall from Definition 22, the characteristic exponent of X is a function,

$$\Psi(\xi) = -\ln \mathbb{E}[e^{i\xi X_1}],$$

which we can calculate using the Lévy-Khintchine formula in Theorem 2.

Theorem 12. *The characteristic exponent for a meromorphic Lévy process, X, with Lévy triple (μ, σ, Π) is*

$$\Psi(\xi) = -i\mu\xi + \frac{1}{2}\sigma^2\xi^2 + \xi^2 \sum_{n\in\mathbb{N}} \left[\frac{a_n}{\rho_n(\rho_n - i\xi)} + \frac{\widehat{a}_n}{\widehat{\rho}_n(\widehat{\rho}_n + i\xi)} \right]. \tag{5.6}$$

Proof. See [36, Theorem 1], where the authors calculate the *Laplace transform*, $\psi(\xi)$, which, by straightforward analytical continuation, has the relationship

$$\Psi(\xi) = -\psi(i\xi).$$

□

Since it is holomorphic everywhere except for an isolated set of poles, $\{-\widehat{\rho}_n, \rho_n\}$, the characteristic exponent of a meromorphic Lévy process is a *meromorphic function* (see Appendix A.1), which motivates the name.

5.1.2 Wiener-Hopf factors

Recall that the Wiener-Hopf factorisation in Theorem 3 states that, for any Lévy process and some $q > 0$,

$$\mathbb{E}\left[e^{i\xi X_{e_q}}\right] = \mathbb{E}\left[e^{i\xi \overline{X}_{\mathbf{e}_q}}\right] \mathbb{E}\left[e^{i\xi \underline{X}_{\mathbf{e}_q}}\right],$$

and the Wiener-Hopf factors are the terms in the product on the right-hand side,

$$\phi^+(\xi) = \mathbb{E}\left[e^{i\xi \overline{X}_{\mathbf{e}_q}}\right], \quad \phi^-(\xi) = \mathbb{E}\left[e^{i\xi \underline{X}_{\mathbf{e}_q}}\right].$$

One of the important properties of meromorphic Lévy processes is that these Wiener-Hopf factors are tractable. More specifically, they can be expressed as convergent infinite products of functions of the roots and poles of the characteristic exponent. It is clear from (5.6) that the characteristic exponent has infinitely many poles, with values given by $\{-\widehat{\rho}_n, \rho_n\}$, and also infinitely many roots, as we verify next.

Theorem 13. *For every $q > 0$, there are infinitely many solutions to the equation $q + \Psi(i\xi) = 0$, which we call $\{-\widehat{\zeta}_n, \zeta_n\}$. These roots are real and interlace with the poles of $\Psi(\xi)$,*

$$\cdots - \widehat{\rho}_2 < -\widehat{\zeta}_2 < -\widehat{\rho}_1 < -\widehat{\zeta}_1 < 0 < \zeta_1 < \rho_1 < \zeta_2 < \rho_2 \cdots$$

Proof. Since $\Psi(\xi)$ is a meromorphic function with $\Psi(0) = 0$, $\Psi(\xi)$ is analytic in some neighbourhood of $0 \in \mathbb{C}$. Then, as stated in the proof of [49, Theorem 2], Ψ can be analytically extended to the complex half-planes $\mathcal{I}(z) > 0$ and $\mathcal{I}(z) < 0$, and we must have

$$\mathcal{I}\left(\frac{-\Psi(i\xi)}{i\xi}\right) > 0, \tag{5.7}$$

for $\xi \in \mathbb{C}$ with $\mathcal{R}(\xi) > 0$, where $\mathcal{R}(z)$ is the real part of z and $\mathcal{I}(z)$ is the imaginary part. Additionally, for $q > 0$ and $\xi \in \mathbb{C}$ with $\mathcal{R}(\xi) > 0$,

$$\mathcal{I}\left(\frac{-q}{i\xi}\right) > 0. \tag{5.8}$$

This is easily shown by writing $i\xi = a + bi$ for $a, b \in \mathbb{R}$, where we must have $b > 0$ to satisfy $\mathcal{R}(\xi) > 0$. Then

$$\frac{-q}{i\xi} = \frac{-q(a - bi)}{(a + bi)(a - bi)} = -\frac{qa}{a^2 + b^2} + \frac{qbi}{a^2 + b^2},$$

such that $\mathcal{I}(-\frac{q}{i\xi}) = \frac{qb}{a^2+b^2}$, which is positive since q and b are positive. Also, for $z_1, z_2 \in \mathbb{C}$, $\mathcal{I}(z_1 + z_2) = \mathcal{I}(z_1) + \mathcal{I}(z_2)$, implying that, if $\mathcal{I}(z_1), \mathcal{I}(z_2) > 0$, $\mathcal{I}(z_1 + z_1) > 0$. Applying this to (5.7) and (5.8), we must have

$$\mathcal{I}\left(\frac{-\Psi(i\xi) - q}{i\xi}\right) > 0.$$

Finally, applying the statement in the proof of [35, Theorem 1], we find that there are infinitely many roots to the equation $q + \Psi(i\xi) = 0$ and these roots interlace with the poles of the meromorphic function, Ψ, as required. □

Finally, the Wiener-Hopf factors of a meromorphic Lévy process are given in Theorem 14, which has the same proof as Theorem 13 since [35, Theorem 1] also provides the form of the Wiener-Hopf factorisation.

Theorem 14. *For $q > 0$, the Wiener-Hopf factors of a meromorphic Lévy process are*

$$\phi_q^+(\xi) = \prod_{n\geq 1} \frac{1 - \frac{i\xi}{\rho_n}}{1 - \frac{i\xi}{\zeta_n}}, \quad \phi_q^-(\xi) = \prod_{n\geq 1} \frac{1 + \frac{i\xi}{\widehat{\rho}_n}}{1 + \frac{i\xi}{\widehat{\zeta}_n}}. \tag{5.9}$$

These products are convergent.

5.1.3 Variance and jump activity

We now want to investigate the path properties of meromorphic Lévy processes, particularly those that are considered in option pricing models. Here, we discuss the path structure, along with the variation and jump activity.

The paths of a meromorphic Lévy process may include a constant drift, a diffusion, and a jump term that decays at an exponential rate, meaning that the paths are càdlàg but do not need to be continuous. Further, the exponential decay of the jump process means that meromorphic processes have fatter tails than a normal distribution.

The class of meromorphic Lévy process includes members that have both infinite and bounded variation and jump activity. As shown in Theorems 15 and 16, both of these properties depend on the choice of parameters.

Theorem 15. *A meromorphic Lévy process, X, with Lévy triple (μ, σ, Π) has bounded variation if and only if*

i. X has no diffusion component: $\sigma = 0$.

ii. The series $\sum_{n\in\mathbb{N}} a_n\rho_n^{-1}$ and $\sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n^{-1}$ converge.

Proof. Suppose that X is a meromorphic Lévy process with Lévy triple (μ, σ, Π) and parameters $\{a_n\}_{n\in\mathbb{N}}$, $\{\widehat{a}_n\}_{n\in\mathbb{N}}$, $\{\rho_n\}_{n\in\mathbb{N}}$, and $\{\widehat{\rho}_n\}_{n\in\mathbb{N}}$ such that

$$\sum_{n\in\mathbb{N}} a_n\rho_n^{-1} < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n^{-1} < \infty.$$

Recall, from Lemma 4, that a Lévy process has bounded variation if and only if $\sigma = 0$ and its Lévy measure satisfies

$$\int_{\mathbb{R}} (1 \wedge |x|)\, \Pi(dx) < \infty.$$

Consider the integral where $1 \wedge |x|$ is replaced by $|x|$. Using the density of the Lévy measure for a meromorphic Lévy process in (5.4),

$$\begin{aligned}
\int_{\mathbb{R}} |x|\, \Pi(dx) &= \int_{\mathbb{R}} |x|\, \nu(x)\, dx \\
&= \int_{\mathbb{R}} |x| \left(\mathbb{1}\{x > 0\} \sum_{n\in\mathbb{N}} a_n\rho_n e^{-\rho_n x} + \mathbb{1}\{x < 0\} \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n e^{\widehat{\rho}_n x} \right) dx \\
&= -\int_{-\infty}^{0} x \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n e^{\widehat{\rho}_n x}\, dx + \int_{0}^{\infty} x \sum_{n\in\mathbb{N}} a_n\rho_n e^{-\rho_n x}\, dx.
\end{aligned}$$

The terms of both summations are non-negative. Thus, by Tonelli's theorem, we can integrate term by term (interchanging the integral and summation):

$$\begin{aligned}
\int_{\mathbb{R}} |x|\,\Pi(dx) &= -\sum_{n\in\mathbb{N}} \int_{-\infty}^{0} x\widehat{a}_n\widehat{\rho}_n e^{\widehat{\rho}_n x}\,dx + \sum_{n\in\mathbb{N}} \int_{0}^{\infty} x a_n \rho_n e^{-\rho_n x}\,dx \\
&= -\sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n \int_{-\infty}^{0} x e^{\widehat{\rho}_n x}\,dx + \sum_{n\in\mathbb{N}} a_n\rho_n \int_{0}^{\infty} x e^{-\rho_n x}\,dx \\
&= -\sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n \left[\frac{(\widehat{\rho}_n x - 1)}{\widehat{\rho}_n^2} e^{\widehat{\rho}_n x}\right]_{-\infty}^{0} + \sum_{n\in\mathbb{N}} a_n\rho_n \left[-\frac{(\rho_n x + 1)}{\rho_n^2} e^{-\rho_n x}\right]_{0}^{\infty} \\
&= \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n \left(\frac{1}{\widehat{\rho}_n^2}\right) + \sum_{n\in\mathbb{N}} a_n\rho_n \left(\frac{1}{\rho_n^2}\right) \\
&= \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n^{-1} + \sum_{n\in\mathbb{N}} a_n\rho_n^{-1}. \qquad (5.10)
\end{aligned}$$

By assumption, both of these series converge. Hence,

$$\int_{\mathbb{R}} |x|\,\Pi(dx) < \infty.$$

Finally, since $(1 \wedge |x|) \leq |x|$ for all $x \in \mathbb{R}$,

$$\int_{\mathbb{R}} (1 \wedge |x|)\,\Pi(dx) \leq \int_{\mathbb{R}} |x|\,\Pi(dx) < \infty;$$

so, by Lemma 4, X has bounded variation.

On the contrary, suppose that at least one of the series diverges:

$$\sum_{n\in\mathbb{N}} a_n\rho_n^{-1} = \infty \quad \text{or} \quad \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n^{-1} = \infty.$$

Then, by (5.10),

$$\int_{\mathbb{R}} |x|\,\Pi(dx) = \sum_{n\in\mathbb{N}} \widehat{a}_n\widehat{\rho}_n^{-1} + \sum_{n\in\mathbb{N}} a_n\rho_n^{-1} = \infty. \qquad (5.11)$$

Recall, from Lemma 13, that meromorphic Lévy measures must satisfy

$$\int_{\mathbb{R}} x^2\,\Pi(dx) < \infty.$$

Since $x^2 \geq |x|$ for all $|x| \geq 1$, this implies that

$$\int_{|x|>1} |x|\,\Pi(dx) \leq \int_{|x|>1} x^2\,\Pi(dx) < \infty.$$

Thus, to satisfy (5.11), we require

$$\int_{|x|<1} |x|\,\Pi(dx) = \infty,$$

which implies that

$$\int_{\mathbb{R}} (1 \wedge |x|)\,\Pi(dx) = \int_{|x|<1} |x|\,\Pi(dx) + \int_{|x|\geq 1} \Pi(dx) = \infty;$$

so, by Lemma 4, X has infinite variation. □

Theorem 16. *A meromorphic Lévy process, X, with Lévy triple (μ, σ, Π) has finite jump activity if and only if the series $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} \widehat{a}_n$ converge.*

Proof. Suppose that X is a meromorphic Lévy process with Lévy triple (μ, σ, Π) and parameters $\{a_n\}_{n\in\mathbb{N}}$, $\{\widehat{a}_n\}_{n\in\mathbb{N}}$, $\{\rho_n\}_{n\in\mathbb{N}}$, and $\{\widehat{\rho}_n\}_{n\in\mathbb{N}}$ such that

$$\sum_{n\in\mathbb{N}} a_n < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \widehat{a}_n < \infty$$

Recall, from Lemma 5, that a Lévy process has finite jump activity if and only if

$$\int_{\mathbb{R}} \Pi(dx) < \infty.$$

Using analogous reasoning to the proof for Theorem 15, we find that

$$\int_{\mathbb{R}} \Pi(dx) = \sum_{n\in\mathbb{N}} \widehat{a}_n + \sum_{n\in\mathbb{N}} a_n.$$

By assumption, both series converge. Hence,

$$\int_{\mathbb{R}} \Pi(dx) < \infty.$$

Thus, by Lemma 5, X has finite jump activity. As with the proof of Theorem 4, the opposite direction follows easily. □

Theorems 15 and 16 can be used to find simple examples of meromorphic Lévy processes with finite or infinite variation and jump activity. These examples will be stated in terms of a_n and ρ_n, but there are identical requirements on $\widehat{a}_n$ and $\widehat{\rho}_n$. The property of finite jump activity depends on the parameters $\{a_n\}_{n\in\mathbb{N}}$, while bounded variation depends on both $\{a_n\}$ and $\{\rho_n\}$, along with the diffusivity, σ. By Theorem 16, any meromorphic process with $a_n = 1$ for all $n \in \mathbb{N}$ has infinite jump activity. Further, using Theorem 15

and the knowledge that the harmonic series diverges, taking $\rho_n = n$ gives an example of a meromorphic process where both variation and jump activity are infinite. A meromorphic process with infinite jump activity but finite variation can be achieved by alternatively taking $\rho_n = n^2$, since we now have a convergent p-series in Theorem 15. Finally, a meromorphic process with parameters $a_n = 1/n^2$ and $\rho_n = n$ has finite variation and jump activity.

These simple examples show that the class of meromorphic Lévy processes admits a wide range of properties in path structure, variation, and jump activity. This would be beneficial in producing option pricing models to fit a particular financial security.

5.2 Examples

To complete this chapter, we define two key subclasses of meromorphic Lévy processes: *β-family* and *hyper-exponential processes*.

5.2.1 β-family processes

The β-family Lévy processes are an important subclass of meromorphic Lévy processes, defined as follows.

Definition 44 (β-family process)**.** *A meromorphic Lévy process, X, belongs to the β-family if its Lévy measure, with form given by* (5.4)*, has parameters:*

$$a_n = \frac{c_1}{\beta_1(\alpha_1 + n - 1)} \binom{n + \lambda_1 - 2}{n - 1}, \quad \rho_n = \beta_1(\alpha_1 + n - 1),$$
$$\widehat{a}_n = \frac{c_2}{\beta_2(\alpha_2 + n - 1)} \binom{n + \lambda_2 - 2}{n - 1}, \quad \widehat{\rho}_n = \beta_2(\alpha_2 + n - 1),$$

for $n \in \mathbb{N}$, where α_i, $\beta_i > 0$, $c_i \geq 0$ and $\lambda_i \in (0, 3)$ for $i = 1, 2$, and $\binom{x}{y} = \frac{x(x-1)\cdots(x-y+1)}{y(y-1)\cdots 1}$ is the combination function.

In this case, the series that define the density of the Lévy measure are convergent, leading to the following representation, which could be used as an alternative definition of the β-family.

Theorem 17. *Let X be a β-family Lévy process with parameters α_i, $\beta_i > 0$, $c_i \geq 0$ and $\lambda_i \in (0, 3)$ for $i = 1, 2$. Then, the Lévy measure, $\Pi(dx)$, has density*

$$\nu(x) = c_1 \frac{e^{-\alpha_1\beta_1 x}}{\left(1 - e^{-\beta_1 x}\right)^{\lambda_1}} \mathbb{1}\{x > 0\} + c_2 \frac{e^{\alpha_2\beta_2 x}}{\left(1 - e^{\beta_2 x}\right)^{\lambda_2}} \mathbb{1}\{x < 0\}.$$

Proof. By (5.4), the Lévy measure of a meromorphic Lévy process has density

$$\nu(x) = \mathbb{1}\{x > 0\} \sum_{n\in\mathbb{N}} a_n \rho_n e^{-\rho_n x} + \mathbb{1}\{x < 0\} \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n e^{\widehat{\rho}_n x}.$$

We shall evaluate the first series; the form of the second series follows by symmetry. First, note that the binomial series expansion of the denominator of the required density is

$$\frac{1}{(1-e^{-\beta_1 x})^{\lambda_1}} = \sum_{k\in\mathbb{Z}_+} \binom{k+\lambda_1-1}{k} \left(e^{-\beta_1 x}\right)^k. \tag{5.12}$$

Substituting the parameters from Definition 44, making a change of variable $n \mapsto k+1$, and then applying (5.12), we have

$$\begin{aligned}
\sum_{n\in\mathbb{N}} a_n \rho_n e^{-\rho_n x} &= \sum_{n\in\mathbb{N}} \frac{c_1}{\beta_1(\alpha_1+n-1)} \binom{n+\lambda_1-2}{n-1} \beta_1(\alpha_1+n-1) e^{-\beta_1(\alpha_1+n-1)x} \\
&= \sum_{n\in\mathbb{N}} c_1 \binom{n+\lambda_1-2}{n-1} e^{-\beta_1(\alpha_1+n-1)x} \\
&= \sum_{k\in\mathbb{Z}_+} c_1 \binom{k+\lambda_1-1}{k} e^{-\beta_1(\alpha_1+k)x} \\
&= c_1 e^{-\alpha_1\beta_1 x} \sum_{k\in\mathbb{Z}_+} \binom{k+\lambda_1-1}{k} \left(e^{-\beta_1 x}\right)^k \\
&= \frac{c_1 e^{-\alpha_1\beta_1 x}}{(1-e^{-\beta_1 x})^{\lambda_1}}.
\end{aligned}$$

By analogous reasoning, we find that

$$\sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n e^{\widehat{\rho}_n x} = c_2 \frac{e^{\alpha_2\beta_2 x}}{\left(1-e^{\beta_2 x}\right)^{\lambda_2}},$$

so that

$$\nu(x) = c_1 \frac{e^{-\alpha_1\beta_1 x}}{\left(1-e^{-\beta_1 x}\right)^{\lambda_1}} \mathbb{1}\{x > 0\} + c_2 \frac{e^{\alpha_2\beta_2 x}}{\left(1-e^{\beta_2 x}\right)^{\lambda_2}} \mathbb{1}\{x < 0\},$$

as required. □

The β-family of meromorphic Lévy processes was introduced by Kuznetsov [34]. One of the beneficial properties of this class is that the characteristic exponent can be expressed in terms of *beta functions* instead of infinite series. This allows increased speed and accuracy in numerical calculations since we are evaluating a known function instead of a truncated infinite series.

Theorem 18. *Let X be a β-family Lévy process with parameters α_i, $\beta_i > 0$, $c_i \geq 0$ and $\lambda_i \in (0,3)$ for $i = 1,2$. Let $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$ be the beta function, and $\psi(x) = \frac{d}{dx}\ln(\Gamma(x))$ be the digamma function. The characteristic exponent of X is*

$$\Psi(\xi) = -i\mu\xi + \frac{1}{2}\sigma^2\xi^2 + J_1(\xi) + J_2(\xi),$$

where $J_1(\xi)$ and $J_2(\xi)$ are the terms contributed by the positive and negative jumps of X, respectively. These terms depend on the values of λ_1 and λ_2:

i. (a) If $\lambda_1 \in (0,3)\backslash\{1,2\}$,

$$\begin{aligned} J_1(\xi) =& \frac{c_1 i\xi}{\beta_1^2} B(\alpha_1,\, 1-\lambda_1)\big(\psi(1+\alpha_1-\lambda_1) - \psi(\alpha_1)\big) \\ &+ \frac{c_1}{\beta_1} B(\alpha_1,\, 1-\lambda_1) - \frac{c_1}{\beta_1} B\left(\alpha_1 - \frac{i\xi}{\beta_1},\, 1-\lambda_1\right). \end{aligned}$$

(b) If $\lambda_2 \in (0,3)\backslash\{1,2\}$,

$$\begin{aligned} J_2(\xi) =& -\frac{c_2 i\xi}{\beta_2^2} B(\alpha_2,\, 1-\lambda_2)\big(\psi(1+\alpha_2-\lambda_2) - \psi(\alpha_2)\big) \\ &+ \frac{c_2}{\beta_2} B(\alpha_2,\, 1-\lambda_2) - \frac{c_2}{\beta_2} B\left(\alpha_2 + \frac{i\xi}{\beta_2},\, 1-\lambda_2\right). \end{aligned}$$

ii. (a) If $\lambda_1 = 1$,

$$J_1(\xi) = \frac{1}{\beta_1}\psi\left(\alpha_1 - \frac{i\xi}{\beta_1}\right) + \frac{1}{\beta_1}\psi(\alpha_1) + \frac{i\xi}{\beta_1^2}\psi'(\alpha_1).$$

(b) If $\lambda_2 = 1$,

$$J_2(\xi) = \frac{1}{\beta_2}\psi\left(\alpha_2 + \frac{i\xi}{\beta_2}\right) + \frac{1}{\beta_2}\psi(\alpha_2) + \frac{i\xi}{\beta_2^2}\psi'(\alpha_2).$$

iii. (a) If $\lambda_1 = 2$,

$$J_1(\xi) = \frac{1}{\beta_1}\left(1 - \alpha_1 + \frac{i\xi}{\beta_1}\right)\left[\psi\left(\alpha_1 - \frac{i\xi}{\beta_1}\right) - \psi(\alpha_1)\right] + \frac{i\xi(1-\alpha_1)}{\beta_1^2}\psi'(\alpha_1).$$

(b) If $\lambda_2 = 2$,

$$J_2(\xi) = \frac{1}{\beta_2}\left(1 - \alpha_2 + \frac{i\xi}{\beta_2}\right)\left[\psi\left(\alpha_2 + \frac{i\xi}{\beta_2}\right) - \psi(\alpha_2)\right] + \frac{i\xi(1-\alpha_2)}{\beta_2^2}\psi'(\alpha_2).$$

Proof. See [34, Proposition 9] □

The β-family itself is a rich class with varied subclasses. The jump distribution is parameterised by λ_i, which controls the small jumps, α_i and β_i, which control the exponential tails, and c_i, which determines the intensity of the jumps. According to [34], some processes in the β-family have similar dynamics to the *KoBoL* processes that were seen in Section 4.1.1 and have been regularly included in option pricing literature. Therefore, it is reasonable to assume that β-family processes will also be useful in this context and their inclusion fills a gap in the literature. As discussed previously, when forming an option pricing model, we are often interested in whether the process has bounded variation or jump activity. The β-family admits paths of both infinite and finite variation and jump activity, providing flexibility to ensure that the assumptions of our models are satisfied.

Theorem 19. *Let X be a β-family Lévy process with parameters α_i, $\beta_i > 0$, $c_i \geq 0$ and $\lambda_i \in (0,3)$ for $i = 1,2$. Then X has bounded variation if and only if $\sigma = 0$ and $\lambda_1, \lambda_2 \in (0,2)$. Further, X has finite jump activity if and only if $\lambda_1, \lambda_2 \in (0,1)$.*

Proof. From Theorem 15, a meromorphic Lévy process with $\sigma = 0$ has bounded variation if and only if

$$\sum_{n\in\mathbb{N}} a_n \rho_n^{-1} + \sum_{n\in\mathbb{N}} a_n \widehat{\rho}_n^{-1} < \infty,$$

or equivalently, since a_n and ρ_n are non-negative,

$$\sum_{n\in\mathbb{N}} a_n \rho_n^{-1} < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n^{-1} < \infty.$$

Consider only the first series; the convergence conditions of the second series are identical (for λ_2 instead of λ_1) by symmetry. Using the parameters from Definition 44 and simple manipulations of the summation,

$$\begin{aligned}
\sum_{n\in\mathbb{N}} a_n \rho_n^{-1} &= \sum_{n\in\mathbb{N}} \frac{c_1}{\beta_1^2(\alpha_1+n-1)^2} \binom{n+\lambda_1-2}{n-1} \\
&= \frac{c_1}{\alpha_1\beta_1^2} + \frac{c_1}{\beta_1^2} \sum_{n\geq 2} \frac{1}{(\alpha_1+n-1)^2} \binom{n+\lambda_1-2}{n-1} \\
&= \frac{c_1}{\alpha_1\beta_1^2} + \frac{c_1}{\beta_1^2} \sum_{n\in\mathbb{N}} \frac{1}{(\alpha_1+n)^2} \binom{n+\lambda_1-1}{n}, \qquad (5.13)
\end{aligned}$$

where $\alpha_1, \beta_1 > 0$, $c_1 \geq 0$, and $\lambda_1 \in (0,3)$.

If $\lambda_1 = 1$, the series in (5.13) converges because $\binom{n}{n} = 1$ for all $n \in \mathbb{N}$, giving a convergent p-series. Thus, X has bounded variation.

If $\lambda_1 \in (0,3)\backslash\{1\}$, let $x_n = \frac{1}{(\alpha_1+n)^2}\binom{n+\lambda_1-1}{n}$, so that we are interested in the convergence of $\sum_{n\in\mathbb{N}} x_n$. Then, x_n has the following series expansion at $n=\infty$,

$$x_n = n^{\lambda_1}\left(\frac{1}{\Gamma(\lambda_1)n^3} + \frac{\lambda_1(\lambda_1-1)-4\alpha_1}{2\Gamma(\lambda_1)n^4} + \mathcal{O}\left(\frac{1}{n^5}\right)\right) = \mathcal{O}\left(\frac{1}{n^{3-\lambda_1}}\right). \tag{5.14}$$

If $\lambda_1 \in [2,3)$, $x_n = \mathcal{O}(1/n^p)$ for $p \in (0,1]$, meaning that the series in (5.13) is a divergent p-series. Thus, X has infinite variation.

Alternatively, if $\lambda_1 \in (0,2)\backslash\{1\}$, let $b_n = n^{\frac{4-\lambda_1}{2}}$. Since $\frac{4-\lambda_1}{2} > 1$, $\sum_{n\in\mathbb{N}} b_n$ is a convergent p-series. Additionally, observe that the exponent in b_n is greater than that in (5.14); since $2-\lambda_1 > 0$,

$$3-\lambda_1 = \frac{6-2\lambda_1}{2} = \frac{4-\lambda_1}{2} + \frac{2-\lambda_1}{2} < \frac{4-\lambda_1}{2}.$$

Thus, we have $x_n = \mathcal{O}\big(n^{-(3-\lambda_1)}\big)$ and $b_n = \mathcal{O}\Big(n^{-\frac{4-\lambda_1}{2}}\Big)$, where $-(3-\lambda_1) > -\frac{4-\lambda_1}{2}$. This means that there exists $N \in \mathbb{N}$ such that, for all $n \geq N$,

$$0 \leq x_n \leq b_n.$$

Since $\sum_{n\in\mathbb{N}} b_n$ converges, so does $\sum_{n\geq N} b_n$. Then, by the comparison test, $\sum_{n\geq N} x_n$ converges. Further, since each x_n is finite, $\sum_{n\in\mathbb{N}} x_n$, the required series in (5.13), must converge also. Thus, X has bounded variation.

We have shown that X has infinite variation if $\lambda_1 \in [2,3)$. This statement would also be true (by symmetry) if $\lambda_2 \in [2,3)$. So, X has bounded variation if $\lambda_1, \lambda_2 \in (0,2)$ and infinite variation if either λ_1 or $\lambda_2 \in [2,3)$.

By analogous reasoning, we can show that X has finite jump activity if and only if $\lambda_1, \lambda_2 \in (0,1)$, where we then have $x_n = \frac{1}{\alpha_1+n}\binom{n+\lambda_1-1}{n} = \mathcal{O}\big(n^{-(2-\lambda_1)}\big)$ and $b_n = n^{\frac{3-\lambda_1}{2}}$. □

5.2.2 Hyper-exponential processes

Another important class of meromorphic Lévy processes is the *hyper-exponential processes*. This class is a generalisation of the Kou model [28], which is a classical option pricing model consisting of the sum of a linear Brownian motion and a compound Poisson process with two-sided exponentially distributed jumps. Hyper-exponential Lévy processes replace the jump distribution with a hyper-exponential distribution. This is equivalent to a truncated meromorphic Lévy process [38].

Definition 45 (Hyper-exponential process)**.** *A meromorphic Lévy process, X, with parameters $\{a_n\}_{n\in\mathbb{N}}$, $\{\widehat{a}_n\}_{n\in\mathbb{N}}$, $\{\rho_n\}_{n\in\mathbb{N}}$, and $\{\widehat{\rho}_n\}_{n\in\mathbb{N}}$ is a hyper-exponential process if there exists $N, \widehat{N} \in \mathbb{N}$ such that $a_{n_1} = \widehat{a}_{n_2} = 0$ for all $n_1 \geq N$ and $n_2 \geq \widehat{N}$.*

We could alternatively interpret a hyper-exponential process as having $\rho_{n_1} = \infty$ and $\widehat{\rho}_{n_2} = \infty$ for each $n_1 \geq N$ and $n_2 \geq \widehat{N}$.

By (5.6), a hyper-exponential process has characteristic exponent

$$\Psi(\xi) = -i\mu\xi + \frac{1}{2}\sigma^2\xi^2 + \xi^2 \sum_{n=1}^{N} \frac{a_n}{\rho_n(\rho_n - i\xi)} + \xi^2 \sum_{n=1}^{\widehat{N}} \frac{\widehat{a}_n}{\widehat{\rho}_n(\widehat{\rho}_n + i\xi)},$$

and, by (5.4), its Lévy measure has density

$$\nu(x) = \mathbb{1}\{x > 0\} \sum_{n=1}^{N} a_n \rho_n e^{-\rho_n x} + \mathbb{1}\{x < 0\} \sum_{n=1}^{\widehat{N}} \widehat{a}_n \widehat{\rho}_n e^{\widehat{\rho}_n x}. \tag{5.15}$$

According to [36], hyper-exponential Lévy processes can achieve the same dynamics as the Kou model, a widely accepted and utilised option pricing method. However, since it is a more general version of the Kou model, an algorithm that uses hyper-exponential Lévy processes will also allow additional freedom in accurately fitting the model to the specific financial security. Accordingly, it is again reasonable to state that the class of hyper-exponential Lévy processes is a useful addition to the option pricing literature.

All results that we have seen for meromorphic Lévy processes also apply to hyper-exponential processes with trivial modifications. In Theorem 13, we now only have a finite number of roots to the equation $q + \Psi(i\xi) = 0$, equal to the number of poles, which includes N positive terms and $\widehat{N}$ negative terms. The Wiener-Hopf factors in Theorem 14 have an identical form, but are products of length N and $\widehat{N}$,

$$\phi_q^+(\xi) = \prod_{n=1}^{N} \frac{1 - \frac{i\xi}{\rho_n}}{1 - \frac{i\xi}{\zeta_n}}, \quad \phi_q^-(\xi) = \prod_{n=1}^{\widehat{N}} \frac{1 + \frac{i\xi}{\widehat{\rho}_n}}{1 + \frac{i\xi}{\widehat{\zeta}_n}}. \tag{5.16}$$

Additionally, the series in Theorems 15 and 16 become finite sums (of finite-valued terms), and hence, must converge. This means that whenever $\sigma = 0$, hyper-exponential Lévy processes must have finite jump activity and bounded variation.

5.2.3 Other examples

There are many other interesting subclasses of meromorphic Lévy processes, such as the *θ-class* and *hyper-geometric* processes, which are presented in [36]. We will not discuss these processes or implement them as drivers in our algorithm, but they all satisfy Theorem 14 and have known semi-explicit Wiener-Hopf factors, so could be used in the *Meromorphic Wiener-Hopf method* that we derive in Chapter 6.

Chapter 6

Meromorphic Wiener-Hopf method

While Lévy-driven option pricing models are currently available, these methods usually must approximate the Wiener-Hopf factors. Such methods have seen abundant success, but each requires a set of specific assumptions and may only apply to a particular subset of Lévy processes. Accordingly, it is always beneficial to introduce additional option pricing algorithms based on different Lévy processes or with different assumptions. Further, if we can find Lévy processes that have explicit expressions for their Wiener-Hopf factors, then the option pricing formula, (4.20), can be immediately evaluated. The meromorphic Lévy processes introduced in Chapter 5 are promising candidates. In this section, we discuss how a barrier option may be modelled using a meromorphic Lévy process and then demonstrate the option pricing formula in this case, calling the final algorithm the *Meromorphic Wiener-Hopf method*. Lastly, we discuss the potential advantages and disadvantages of this technique.

6.1 Numerical implementation of meromorphic processes

To define an algorithm driven by meromorphic Lévy processes, we need to be able to numerically evaluate the characteristic exponent and Wiener-Hopf factors in Theorems 12 and 14, respectively. While meromorphic Lévy processes do have known expressions for these functions, they involve infinite series and products that must be truncated for numerical evaluation, which is reasonable since, as discussed in Chapter 5, they are convergent. However, there is currently no convention for how many terms are required since it is dependent on the parameters of the process. Instead, we determine if the process has approximately converged by trial and error; if a significant increase in terms leads to an insignificant change in the result, we assume convergence. Additionally, recall that if convergence is not achieved, we have a hyper-exponential process and, as seen in Section 5.2.2, all of the formulae are identical. Given two truncation points, N and $\widehat{N}$, for each

of the two series/products, the characteristic exponent of a meromorphic Lévy process is approximated by

$$\Psi(\xi) \approx -i\mu\xi + \frac{1}{2}\sigma^2\xi^2 + \xi^2 \sum_{n=1}^{N} \frac{a_n}{\rho_n(\rho_n - i\xi)} + \xi^2 \sum_{n=1}^{\widehat{N}} \frac{\widehat{a}_n}{\widehat{\rho}_n(\widehat{\rho}_n + i\xi)},$$

and similar for the density of the Lévy measure and Wiener-Hopf factors.

The roots $\{-\widehat{\zeta}_n, \zeta_n\}_{n\in\mathbb{N}}$ of the characteristic exponent also need to be calculated. Theorem 13 states that the roots interlace with the poles of the process, such that, for $n \geq 2$,

$$\zeta_1 \in (0, \rho_1), \quad \widehat{\zeta}_1 \in (0, \widehat{\rho}_1), \quad \zeta_n \in (\rho_{n-1},\, \rho_n), \quad \widehat{\zeta}_n \in (\widehat{\rho}_{n-1},\, \widehat{\rho}_n).$$

Then, the interval bisection method can be used to approximate the root in each of these intervals (see Appendix A.3). This technique was utilised for some particular examples of meromorphic Lévy processes in [34], where they state that this procedure is highly efficient and could be computed in parallel for further improvements in speed.

Now that we have techniques for truncating the infinite product and then calculating the roots in (5.9), we are able to implement the option pricing formula in (4.20) for meromorphic Lévy processes. The last consideration in our algorithm is the regions on which the formulae will be evaluated. Recall that the underlying security price is given by $S_t = S_0 e^{X_t}$, where S_0 is the initial value at the start of the contract. Of course, we want the final output of our algorithm to initiate at S_0. Thus, the space region, where the Lévy process is evaluated, is given by an interval centered around 0. In [32], this is called the *localisation region* and is defined to be $[-L\ln(2), L\ln(2))$, where $L \in \mathbb{N}$ is a scaling factor which is the smallest value such that having a larger domain does not significantly alter the resultant output, chosen by trial and error. This scaling factor is usually a low value and is relatively easy to find; we use $L = 1$ in all of our experiments. The step size in this domain, $s \in \mathbb{R}_+$, is a parameter of the algorithm, again chosen by trial and error, so that increasing the step size does not affect the output. We use a step size of $s = 0.001$ in our experiments. Finally, after the Fourier transform is applied, the data is in a new domain, called the *frequency domain*, which is given by the interval $[-\frac{\pi}{s},\, s)$ [32].

6.2 The Algorithm

Finally, we derive our new option pricing method. In Chapter 3, we discussed various reasonable assumptions of option pricing models, such as the presence of large jumps, fat-tailed distributions, and the concept of arbitrage. Chapter 4 showed that these requirements could be effectively satisfied by exponential Lévy processes. We demonstrated how the Wiener-Hopf factorisation could be used to derive an equation for the option price in this situation, then discussed several previous models which implemented these ideas and showed their accuracy. Many of these algorithms aimed to approximate the

Wiener-Hopf factors. However, in Chapter 5, meromorphic Lévy processes were defined and shown to be a rich class of processes that have the required properties, such as non-continuous paths and exponentially decaying tails. Additionally, this class has a known, semi-explicit formula for its Wiener-Hopf factors, which was shown to be numerically tractable in the previous section. Thus, we can now define an option pricing method that is driven by meromorphic Lévy processes, using the formula in (4.20) with the Wiener-Hopf factors in (5.9). We call our proposed meromorphic-driven option pricing algorithm the *Meromorphic Wiener-Hopf* (MWH) method. It can be summarised as follows.

Preliminary Steps:

Consider a knock-out barrier put option that depends on the value of a financial security which is modelled according to a given exponential meromorphic Lévy process. Choose an equivalent martingale measure (EMM) based on data from the given security.

Step 1: Input the risk-free interest rate of the market, r, the initial security spot price, S_0, and the option parameters: strike price, K, barrier, H, and time until maturity, T.

Step 2: Choose a meromorphic Lévy process to model the log-security price. Input the parameters σ, $\{a_n\}_{n\in\mathbb{N}}$, $\{\rho_n\}_{n\in\mathbb{N}}$, $\{\widehat{a}_n\}_{n\in\mathbb{N}}$, $\{\widehat{\rho}_n\}_{n\in\mathbb{N}}$. Choose two truncation points for the process: N, corresponding to $\{a_n\}_{n\in\mathbb{N}}$, $\{\rho_n\}_{n\in\mathbb{N}}$, and $\widehat{N}$, corresponding to $\{\widehat{a}_n\}_{n\in\mathbb{N}}$, $\{\widehat{\rho}_n\}_{n\in\mathbb{N}}$.

Step 3: Calculate the characteristic exponent of the Lévy process using (5.6),

$$\Psi(\xi) = -i\mu\xi + \frac{1}{2}\sigma^2\xi^2 + \xi^2\sum_{n=1}^{N}\frac{a_n}{\rho_n(\rho_n - i\xi)} + \xi^2\sum_{n=1}^{\widehat{N}}\frac{\widehat{a}_n}{\widehat{\rho}_n(\widehat{\rho}_n + i\xi)},$$

the density of the Lévy measure using (5.15),

$$\nu(x) = \mathbb{1}\{x > 0\}\sum_{n=1}^{N} a_n\rho_n e^{-\rho_n x} + \mathbb{1}\{x < 0\}\sum_{n=1}^{\widehat{N}}\widehat{a}_n\widehat{\rho}_n e^{\widehat{\rho}_n x},$$

and the drift μ using (4.2),

$$\mu = r + \frac{\sigma^2}{2} + \int_{\mathbb{R}}\left(1 - e^x + x\mathbb{1}\{|x| < 1\}\right)\Pi(dx).$$

Step 4: Choose the space step size, s, the number of Post-Widder samples, N_{pw}, and the acceleration constant, κ, if the Post-Widder acceleration technique is required. Convergence of the Post-Widder formula usually requires $N_{pw} \approx 500$ without acceleration, or $N_{pw} \approx 5$, $\kappa = 3$ with acceleration.

Step 5: Set $h = \ln\left(\frac{H}{S_0}\right)$ and $q = \frac{N_{pw}}{T}$.

Step 6: Choose a scaling factor $L \in \mathbb{N}$ and set the localisation domain to be $[-L\ln(2),\, L\ln(2))$.

Step 7: Set the number of space steps to be $M = 2^m$ for $m \in \mathbb{N}$ such that $2^{m-1} < \frac{2L\ln(2)}{s} \leq 2^m$.

Step 8: Calculate the space grid points,

$$x = -\frac{1}{2}Ms + ks \quad \text{for } k = 0, 1, \ldots, M-1,$$

and the frequency domain grid points,

$$\xi = \frac{2\pi\ell}{Ms} \quad \text{for } \ell = -\frac{M}{2}, \ldots, -1, 0, 1, \ldots, \frac{M}{2} - 1.$$

So, the space domain is $\left[-\frac{1}{2}Ms,\, \frac{1}{2}Ms\right)$ and the frequency domain is $\left[-\frac{\pi}{s},\, \frac{\pi}{s}\right)$.

Barrier Option Pricing

Step 1: For each space grid point, x, calculate $v_0(N_{pw}/T,\, x) = \widetilde{G}(x) = (K - S_0 e^x)^+$.

Step 2: Calculate the roots, $\{-\widehat{\zeta}_n,\, \zeta_n\}$ of the equation $q + r + \Psi(i\xi) = 0$ using the bisection method with intervals

$$\begin{aligned} \zeta_1 &\in (0, \rho_1), \quad \zeta_n \in (\rho_{n-1},\, \rho_n) \quad n \leq N, \\ \widehat{\zeta}_1 &\in (0, \widehat{\rho}_1), \quad \widehat{\zeta}_n \in (\widehat{\rho}_{n-1},\, \widehat{\rho}_n) \quad n \leq \widehat{N}. \end{aligned}$$

Step 3: For each ξ, calculate the Wiener-Hopf factors using (5.16),

$$\phi_q^+(\xi) = \prod_{n=1}^{N} \frac{1 - \frac{i\xi}{\rho_n}}{1 - \frac{i\xi}{\zeta_n}}, \qquad \phi_q^-(\xi) = \prod_{n=1}^{\widehat{N}} \frac{1 + \frac{i\xi}{\widehat{\rho}_n}}{1 + \frac{i\xi}{\widehat{\zeta}_n}}.$$

Step 4: For $p = 1, 2, \ldots, N_{pw}$, calculate $v_p(q, x)$ for each x, according to (4.20),

$$\begin{aligned} v_0(q, x) &= (K - S_0 e^x)^+ \, \mathbb{1}\{x > h\}, \\ v_p(q, x) &= \frac{q}{q+r} \, \mathcal{F}^{-1} \, \phi_{q+r}^-(\xi) \, \mathcal{F} \, \mathbb{1}_{(h,\infty)} \, \mathcal{F}^{-1} \, \phi_{q+r}^+(\xi) \, \mathcal{F} \, V_{p-1}(q, x), \quad \text{for } p \in \mathbb{N}. \end{aligned}$$

Step 5: The estimated option value is $V(T, x) = v_{N_{pw}}(q, 0)$.

Our Meromorphic Wiener-Hopf algorithm can be implemented in Python as follows. Consider, again, a knock-out barrier put option that depends on a financial security that is modelled according to a given exponential meromorphic Lévy process. Suppose that the meromorphic Lévy process can be truncated to have N positive roots and $\widehat{N}$ negative roots, so that the parameters of this known process, **rho**, **a**, **rho_hat**, and **a_hat**, can be represented as vectors of length N and $\widehat{N}$. These known parameters, along with the parameters of the option and the Wiener-Hopf algorithm are input. The space and frequency domain values are also input.

```
rho = ...
rho_hat = ...
a = ...
a_hat = ...

# --- Option constants --- #
K = ...      # strike price
S0 = ...     # spot price
H = ...      # barrier price
r = ...      # risk-free interest rate
T = ...      # Maturity (years)

# --- Wiener-Hopf algorithm constants --- #
h = np.log(H / S0)  # log-barrier price
N_pw = ...          # Number of post-widder samples
q = N_pw / T        # Observation rate

M = ...             # Number of space steps
xstep = ...         # Space step size
x = [-M * xstep/2 + k * xstep for k in range(0, M+1)]  # Space ticks
xi = np.zeros(M+1)  # Frequency ticks
for l in range(-int(M / 2), int(M / 2)):
    xi[l] = 2 * pi * l/(xstep * M)
```

Using the given parameters, the required μ value can be calculated as follows.

```
def calc_mu(sigma, a, rho, a_hat, rho_hat, r, dividend):
    def mero_integrand(x, a, rho, a_hat, rho_hat):
        if x > 0:
            density = sum(np.multiply(a, np.multiply(
                        rho, np.exp(-np.multiply(rho, x)))))
        elif x < 0:
            density = sum(np.multiply(a_hat, np.multiply(
                        rho_hat, np.exp(np.multiply(rho_hat, x)))))
        else:
            return 0
        indicatorfunc = 0
```

```
        if (-1 < x) & (x < 1):
            indicatorfunc = 1
        return (1 - np.exp(x) + x*indicatorfunc)*density

    mu = r - dividend + sigma**2/2 +
         quad(mero_integrand, -np.inf, np.inf,
              args=(a, rho, a_hat, rho_hat))[0]
    return mu

mu = calc_mu(sigma, a, rho, a_hat, rho_hat, r, 0)
```

Next, the roots of the equation $q+r+\Psi(i\xi) = 0$ are calculated using the bisection method, implemented using the function **bisection** that is defined in Appendix A.3. These roots are held in the arrays **zeta** and **zeta_hat**. In these calculations, the characteristic exponent evaluated at $i\xi$ is implemented as **WH_func(xi)**.

```
def WH_func(xi):
    Psi = mu*xi - 1/2 * sigma**2 * xi**2 -
            xi**2 * np.dot(a, 1/np.multiply(p, p + xi)) -
            xi**2 * np.dot(a_hat, 1/np.multiply(p_hat, p_hat - xi))
    return q + r + Psi

zeta = dict.fromkeys(range(N))
zeta[0] = bisection(WH_func, 0, p[0], eps, iter)
for i in range(1, N):
    zeta_plus[i] = bisection(WH_func, p[i-1], p[i], eps, iter)

zeta_hat = dict.fromkeys(range(N_hat))
zeta_hat[0] = -bisection(WH_func, -p_hat[0], 0, eps, iter)
for i in range(1, N_hat):
    zeta_hat[i] = -bisection(WH_func, -p_hat[i], -p_hat[i-1], eps, iter)
```

The calculated roots can be verified visually by plotting the characteristic exponent along with the values of **zeta** and **zeta_hat**. This is shown in Figure 6.1 for an example with $q + r = 1$, but we could have chosen any $q + r > 0$. Observe that the calculated values coincide with the roots of the equation $q + r + \Psi(i\xi) = 0$, as expected.

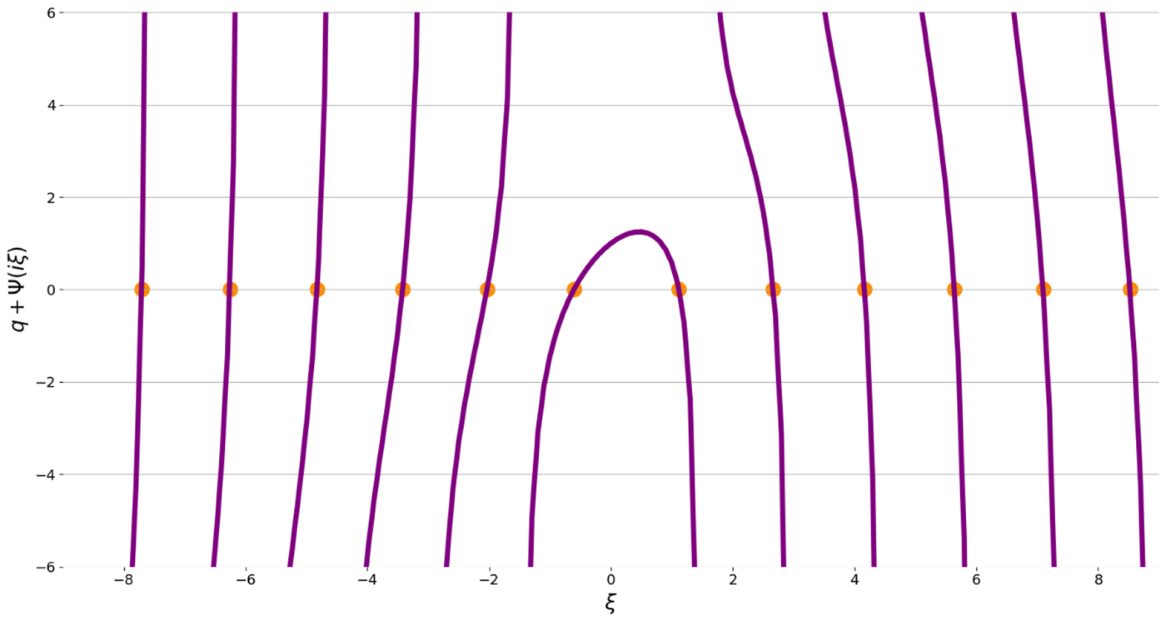


Figure 6.1: Plot of roots of $q+\Psi(i\xi) = 0$ for a β-family Lévy process with $q = 1$, calculated using the bisection method.

The Wiener-Hopf factors for the given meromorphic Lévy process can then be calculated according to their known expressions. These functions are evaluated for each term in the frequency domain and stored in arrays **phi_plus** and **phi_minus**.

```
def WH_factor_plus(x):
    return np.prod([(1 - 1j*x/p[n])/(1 - 1j*x/zeta_plus[n])
                    for n in range(N)])

def WH_factor_minus(x):
    return np.prod([(1 + 1j*x/p_hat[n])/(1 + 1j*x/zeta_minus[n])
                    for n in range(N_hat)])

phi_plus = np.zeros(len(xi), dtype=complex)
phi_minus = np.zeros(len(xi), dtype=complex)
for ixi in range(len(xi)):
    phi_plus[ixi] = WH_factor_plus(xi[ixi])
    phi_minus[ixi] = WH_factor_minus(xi[ixi])
```

Finally, the initial term, $v_0(T, x)$, is calculated for each value in the space domain, and then the iterative formula is used to return the estimate of the option value. This finalises the implementation of the MWH method.

```
def calc_G(K, S0, x):
    return max(K - S0 * np.exp(x), 0)

v = np.zeros((N_pw, len(x)))
for ix in range(len(x)):
    v[0, ix] = calc_G(K, S0, x[ix])

indicatorvec = np.zeros(len(x))
for i in range(len(indicatorvec)):
    if x[i] > h:
        indicatorvec[i] = 1

for i in range(1, N_pw):
     v[i] = (1 / (1 + r / q)) * \
            ifft(np.multiply(phivals_minus, fft(
            np.multiply(indicatorvec,
            ifft(np.multiply(phivals_plus, fft(v[i - 1, :])))))))

# Output option value estimate from final iteration
V = v[-1, :]
V_out = V[min(enumerate(np.exp(x)), key=lambda x: abs(x[1] - 1))[0]]
```

6.3 Simple Wiener-Hopf method under meromorphic processes

Our MWH method has the same goals as the Simple Wiener-Hopf (SWH) method which was discussed in Section 4.5. The SWH method may be implemented for certain subclasses of meromorphic Lévy processes that satisfy the required assumptions. Recall, from Section 4.4, that the underlying process must have no diffusion component and jumps of finite activity, which we know is possible for meromorphic Lévy processes due to Theorem 16. Within this subclass of meromorphic processes, the SWH and MWH methods should produce similar results. In this section, we discuss the implementation of the SWH method for these meromorphic Lévy processes, which will later allow us to compare our models.

Suppose that the underlying Lévy process, $X = \{X_t : t \in \mathbb{R}_+\}$, is in the meromorphic class, with drift μ, diffusion $\sigma = 0$, and parameters $\{a_n\}$, $\{\rho_n\}$, $\{\widehat{a}_n\}$, and $\{\widehat{\rho}_n\}$ such that $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} \widehat{a}_n$ converge. By Theorem 16, this process has finite jump activity and satisfies the requirements of the SWH method. Therefore, there are two subordinators, $X^+ = \{X_t^+ : t \in \mathbb{R}_+\}$ and $-X^- = \{-X_t^- : t \in \mathbb{R}_+\}$, such that

$$X_t = X_t^+ - (-X_t^-).$$

Let the characteristic exponents of X^+ and X^- be Ψ^+ and Ψ^-, respectively. Recall that the meromorphic Lévy process has Lévy measure $\Pi(dx)$, with density

$$\nu(x) = \mathbb{1}\{x > 0\} \sum_{n\in\mathbb{N}} a_n \rho_n e^{-\rho_n x} + \mathbb{1}\{x < 0\} \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n e^{\widehat{\rho}_n x}.$$

Clearly, the density of the positive jumps is given by

$$\nu_+(x) = \mathbb{1}\{x > 0\} \sum_{n\in\mathbb{N}} a_n \rho_n e^{-\rho_n x},$$

and the negative jumps have

$$\nu_-(x) = \mathbb{1}\{x < 0\} \sum_{n\in\mathbb{N}} \widehat{a}_n \widehat{\rho}_n e^{\widehat{\rho}_n x}.$$

Accordingly, the two subordinators, X^+ and $-X^-$, are defined with these jump densities. As with the KoBoL process in Section 4.4.1, the drift term is added to one of the subordinators depending on its sign. If $\mu > 0$,

$$\Psi_+(\xi) = -i\mu\xi + \int_{\mathbb{R}} \left(1 - e^{ix\xi} + ix\xi \mathbb{1}_{\{|x|<1\}}\right) \nu_+(x)\, dx = -i\mu\xi + \xi^2 \sum_{n\in\mathbb{N}} \frac{a_n}{\rho_n(\rho_n - i\xi)},$$
$$\Psi_-(\xi) = \int_{\mathbb{R}} \left(1 - e^{ix\xi} + ix\xi \mathbb{1}_{\{|x|<1\}}\right) \nu_-(x)\, dx = \xi^2 \sum_{n\in\mathbb{N}} \frac{\widehat{a}_n}{\widehat{\rho}_n(\widehat{\rho}_n + i\xi)},$$

and if $\mu < 0$,

$$\Psi_+(\xi) = \int_{\mathbb{R}} \left(1 - e^{ix\xi} + ix\xi \mathbb{1}_{\{|x|<1\}}\right) \nu_+(x)\, dx = \xi^2 \sum_{n\in\mathbb{N}} \frac{a_n}{\rho_n(\rho_n - i\xi)},$$
$$\Psi_-(\xi) = -i\mu\xi + \int_{\mathbb{R}} \left(1 - e^{ix\xi} + ix\xi \mathbb{1}_{\{|x|<1\}}\right) \nu_-(x)\, dx = -i\mu\xi + \xi^2 \sum_{n\in\mathbb{N}} \frac{\widehat{a}_n}{\widehat{\rho}_n(\widehat{\rho}_n + i\xi)}.$$

Then, the SWH method can applied in the same way that was demonstrated in Section 4.4.1. Further, observe that X^+ and X^- are meromorphic Lévy processes with $\widehat{a}_n = 0$ and $a_n = 0$ for all $n \in \mathbb{N}$, respectively. Thus, the WHMC technique from Section 6.4 can be applied to simulate these processes, as shown in Figure 6.2. Observe that X^+ has a drift component, while X^- is a pure jump process.

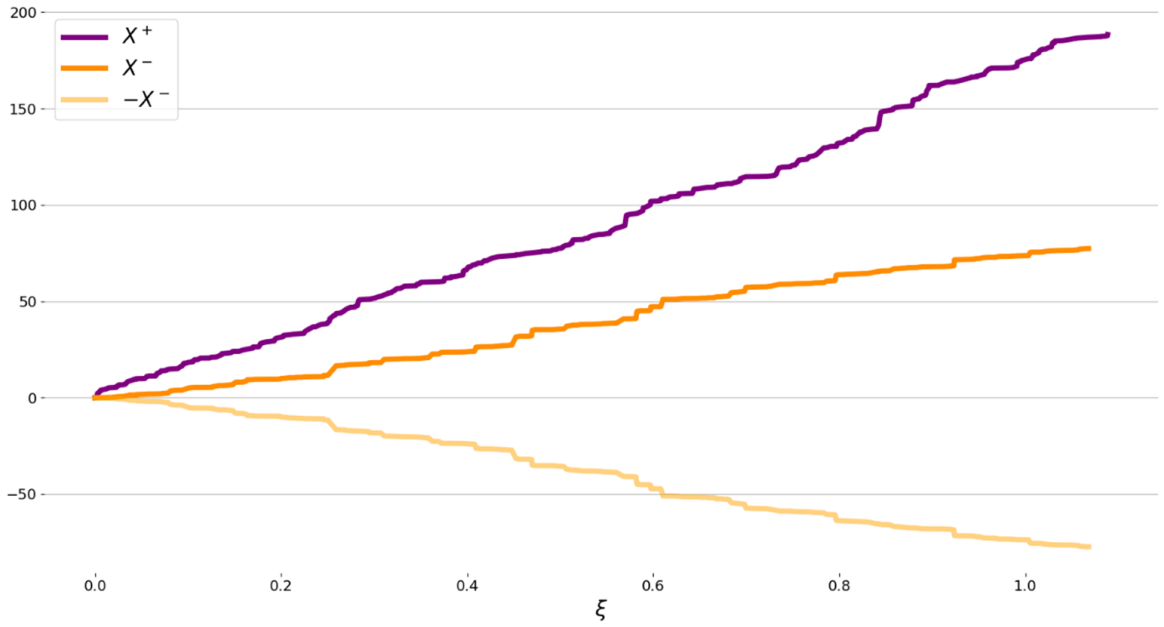


Figure 6.2: Plot of subordinators, X^+ and X^- that decompose a meromorphic Lévy process.

6.4 Meromorphic Wiener-Hopf Monte Carlo simulation

As a comparison for our MWH method, we also want to implement the Wiener-Hopf Monte Carlo (WHMC) simulation method (Section 4.5) for meromorphic Lévy processes. This simulation technique is summarised in Theorem 10, where the joint distribution of a given Lévy process, $X = \{X_t : t \in \mathbb{R}_+\}$, and its supremum, up to time T with M steps, are approximated by variables which are easily calculated if we can simulate random variables with distribution $\overline{X}_{\mathbf{e}_1 T/M}$ and $\underline{X}_{\mathbf{e}_1 T/M}$.

Following Theorem 10, we want to define $\{S_j^{(M/T)}\}$ and $\{I_j^{(M/T)}\}$ with $S_0^{(M/T)} = I_0^{(M/T)} = 0$, $S_j^{(M/T)} \overset{d}{=} \overline{X}_{\mathbf{e}_1 T/M}$, and $I_j^{(M/T)} \overset{d}{=} \underline{X}_{\mathbf{e}_1 T/M}$ for $j = 1, 2, \ldots, M$, independently. For any meromorphic Lévy process, we can calculate the density of these random variables according to [36, Lemma 1, Theorem 2], which gives us Theorem 20.

Theorem 20. *Consider a Lévy process $X = \{X_t : t \in \mathbb{R}_+\}$. For $n \in \mathbb{N}$, define*

$$u_0(\rho, \zeta) := \prod_{k \in \mathbb{N}} \frac{\zeta_k}{\rho_k}, \qquad u_n(\rho, \zeta) := \left(1 - \frac{\zeta_n}{\rho_n}\right) \prod_{\substack{k>1 \\ k \neq n}} \frac{1 - \frac{\zeta_n}{\rho_k}}{1 - \frac{\zeta_n}{\zeta_k}},$$

and

$$\overline{u}(\rho,\zeta) := \left[u_0(\rho,\zeta),\, u_1(\rho,\zeta),\, u_2(\rho,\zeta),\, \dots\,\right]^\top,$$
$$\overline{v}(\zeta,x) := \left[\delta_0(x),\, \zeta_1 e^{-\zeta_1 x},\, \zeta_2 e^{-\zeta_2 x},\, \dots\,\right]^\top.$$

Then, the densities of $\overline{X}_{\mathbf{e}_q}$ *and* $-\underline{X}_{\mathbf{e}_q}$ *are*

$$\mathbb{P}(\overline{X}_{\mathbf{e}_q} \in dx) = \overline{u}(\rho,\zeta)^\top\, \overline{v}(\zeta,x)\, dx, \quad \mathbb{P}(-\underline{X}_{\mathbf{e}_q} \in dx) = \overline{u}(\widehat{\rho},\widehat{\zeta})^\top\, \overline{v}(\widehat{\zeta},x)\, dx,$$

respectively.

For a given meromorphic Lévy process, suppose that we have a set number of positive and negative roots, N and $\widehat{N}$, respectively (*i.e.*, the variables below are given arrays where **rho** and **a** have length N and **rho_hat** and **a_hat** have length $\widehat{N}$).

```
rho = ...
rho_hat = ...
a = ...
a_hat = ...
```

First, we calculate the roots of the equation $q + r + \Psi(i\xi) = 0$ using the same techniques as in Section 6.2. In the code, these variables are contained in arrays of length N and $\widehat{N}$, called **zeta** and **zeta_hat**. Then, define the probability density functions of S and I,

$$f_S = \mathbb{P}(\overline{X}_{\mathbf{e}_q} \in dx), \quad f_I = \mathbb{P}(\underline{X}_{\mathbf{e}_q} \in dx).$$

According to Theorem 20, we have known expressions for these densities for any meromorphic Lévy process. The functions $\overline{u}(\rho,\zeta)$ and $\overline{u}(\widehat{\rho},\widehat{\zeta})$ are represented by arrays **u_plus** and **u_minus**. Since $\overline{u}$ in Theorem 20 does not depend on x, we represent this function as a vectors of its outputs for each possible input value. Alternatively, $\overline{v}(\zeta,x)$ is defined as a function of an array, **zeta** and a value **x**.

```
u_plus = dict.fromkeys(range(N))
u_minus = dict.fromkeys(range(N_hat))

u_plus[0] = np.prod([(zeta[k] / rho[k]) for k in range(N)])
for n in range(1, N):
    u_plus[n] = (1 - zeta[n] / rho[n]) *
    np.prod([(1 - zeta[n] / rho[k]) / (1 - zeta[n] / zeta[k])
             for k in range(1, N) if k != n])

u_minus[0] = np.prod([(zeta_hat[k] / rho_hat[k])
                      for k in range(N_hat)])
```

```
for n in range(1, N_hat):
    u_minus[n] = (1 - zeta_hat[n] / rho_hat[n]) *
                 np.prod([(1 - zeta_hat[n] / rho_hat[k]) /
                          (1 - zeta_hat[n] / zeta_hat[k])
                          for k in range(1, N_hat) if k != n])

def v(zeta, x):
    return np.concatenate(([1 if x == 0 else 0],
                           [zeta_plus[k]*np.exp(-zeta_plus[k]*x)
                            for k in range(len(zeta))])
```

Then, the required densities are given by the following functions.

```
    def f_S(x):
        return np.multiply(u_plus, v(zeta, x))

    def f_I(x):
        return np.multiply(u_minus, v(zeta_hat, x))
```

Given the probability density functions, f_S and f_I, we apply an acceptance-rejection sampling technique (see Appendix A.4) to sample from S and I. Observe that these densities are the scaled summations of N and $\widehat{N}$ exponential densities with increasing rates, which behave similarly to an individual exponential distribution. For large x, f_S and f_I are dominated by the corresponding exponential with the smallest rate, $g_S = \zeta_1 e^{-\zeta_1 x}$ and $g_I = \widehat{\zeta}_1 e^{-\widehat{\zeta}_1 x}$; the sum of exponentials decreases at a greater rate than the single exponential. Scaling g_S and g_I by the maximum value of our densities, which occurs at $x = 0$, we have the scaled functions $g_S^* = \frac{f_S(0)}{g_S(0)} f_S(x)$ and $g_I^* = \frac{f_I(0)}{g_I(0)} f_S(x)$. For all $x \in \mathbb{R}_+$,

$$g_S^*(x) \geq f_S(x), \quad g_I^*(x) \geq g_I(x).$$

These functions are evaluated over the interval $x \in (0, 10]$ in Figure 6.3 to demonstrate this conclusion. This particular example is given with parameters ρ and a that define a β-family process (Section 5.2.1), but applies identically to any meromorphic Lévy process. The functions g_S, g_S^*, g_I, and g_I^* are implemented as **g_S**, **g_S2**, **g_I**, **g_I2**, respectively.

```
def g_S(x):
    return (-zeta_minus[0]) * np.exp(zeta_minus[0] * x)

def g_S2(x):
    return f_S(0)/g_S(0) * g_S(x)

def g_I(x):
    return (zeta_plus[0]) * np.exp(-zeta_plus[0] * x)

def g_I2(x):
    return f_I(0)/g_I(0) * g_I(x)
```

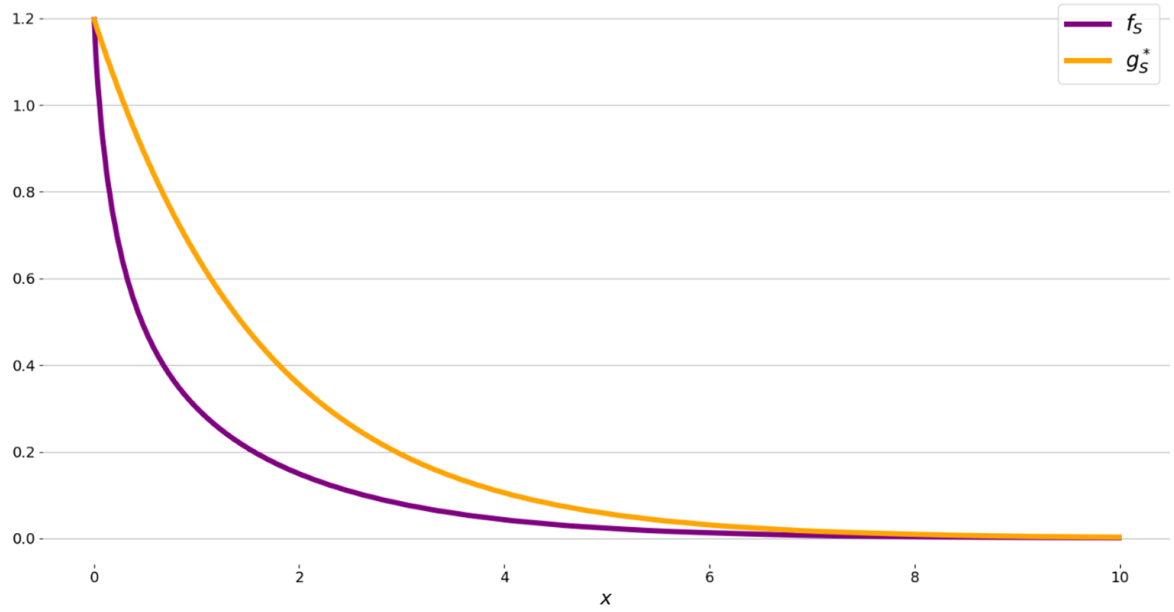


Figure 6.3: Probability density of $S_j = \overline{X}_{\mathbf{e}_1 T/M}$, $f_S(x)$, and a dominating function $g_S^*(x)$ over $x \in [0, 10]$ for a β-family Lévy process with $\mu = 1$, $\sigma = 0.5$, $\alpha_1 = \alpha_2 = 1$, $\beta_1 = \beta_2 = 1.5$, $\lambda_1 = \lambda_2 = 1.5$, $c_1 = c_2 = 1$ and $N = \widehat{N} = 1000$. The observation rate for the roots is $q = M/T = 500/2$.

Thus, we can apply rejection sampling on S and I using $g_S^*(x)$ and $g_I^*(x)$. This allows us to calculate $A_j^{(M/T)}$ and $B_j^{(M/T)}$, as defined in Theorem 10. For $j = 1, 2, \ldots, M$,

$$A_j^{(M/T)} = A_{j-1}^{(M/T)} + S_j^{(M/T)} + I_j^{(M/T)},$$
$$B_j^{(M/T)} = \max\left(B_{j-1}^{(M/T)},\, A_{j-1}^{(M/T)} + S_j^{(M/T)}\right),$$

and $A_0^{(M/T)} = B_0^{(M/T)} = 0$. These variables approximate X and $\overline{X}$ on a random grid.

```
X = np.zeros(M)
X_sup = np.zeros(M)
g = np.zeros(M)
for i in range(1, M):
    g[i] = g[i - 1] + (T / M) * np.random.exponential(1)

def calc_AR_S_ratio(x):
    return f_S(x) / (g_S2(x))

def calc_AR_I_ratio(x):
    return f_I(x) / (g_I2(x))
```

```
I_AR = np.zeros(M)
S_AR = np.zeros(M)
for i in range(1, len(S_AR)):
    accepted = False
    while not (accepted):
        Y = - np.log(np.random.uniform()) / (-zeta_minus[0])
        U = np.random.uniform()
        if U <= calc_AR_S_ratio(Y):
            S_AR[i] = Y
            accepted = True

for i in range(1, len(I_AR)):
    accepted = False
    while not (accepted):
        Y = - np.log(np.random.uniform()) / zeta_plus[0]
        U = np.random.uniform()
        if U <= calc_AR_I_ratio(Y):
            I_AR[i] = -Y
            accepted = True

A = np.zeros(M)
B = np.zeros(M)
for i in range(1, M):
    A[i] = A[i - 1] + S_AR[i] + I_AR[i]
    B[i] = max(B[i - 1], A[i - 1] + S_AR[i])

X = V
X_sup = J
```

The WHMC technique allows us to simulate the path of a given meromorphic Lévy process at random points defined by the arrivals of a Poisson process that converge to the prescribed time frame as the number of steps increases. This algorithm is demonstrated with five independent realisations of a β-family Lévy process in Figure 6.4, where the paths between the random points have been included for visualisation. Monte Carlo estimates of the option price are obtained by simulating a large number of realisations of the required process, then calculating and averaging the payout of the corresponding option. This will be employed for the numerical examples in Section 7.

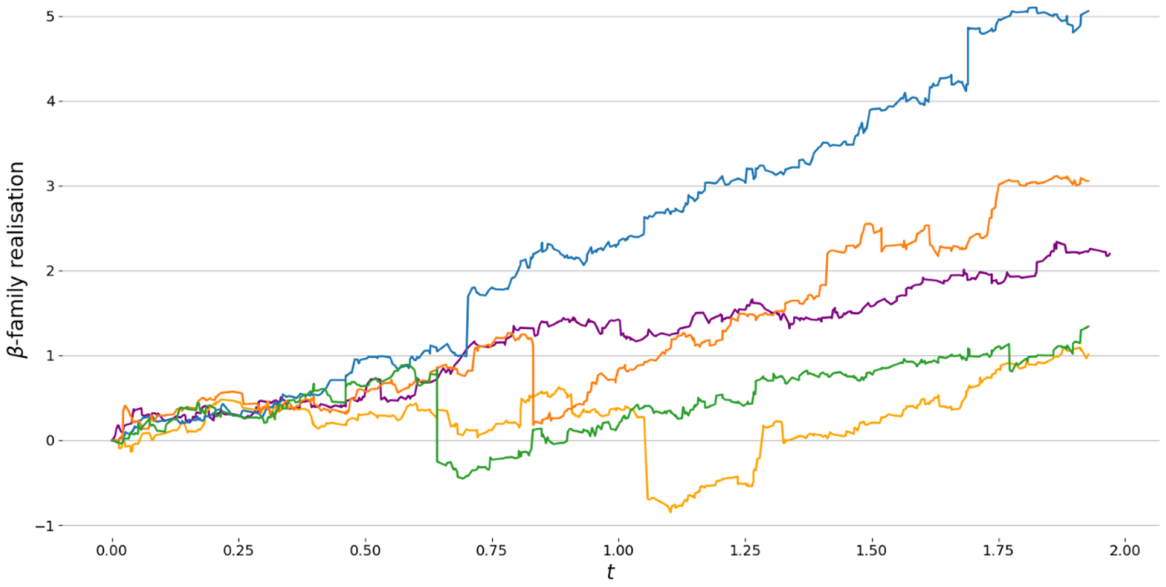


Figure 6.4: Plot of five independent realisations of a β-family Lévy process using the WHMC techniques for meromorphic Lévy processes.

6.5 Evaluation of the MWH method

Now that we have derived a new option pricing method, we evaluate the accuracy, speed, and versatility. Due to the similarity, the SWH method is used for comparison in these criteria.

As with any numerical algorithm, it is crucial that our option pricing method accurately estimates the true value. The SWH method was shown to have high accuracy for a wide range of processes in [33]. Due to the similarity of the MWH method, we would expect this accuracy to be preserved, which is demonstrated to be the case in Chapter 7. Further, since a semi-explicit expression is used for the Wiener-Hopf factors instead of an approximation, the accuracy may be improved in some situations. However, contrarily, if the truncation of the infinite products that we saw in (5.9) is not sufficient, then the semi-explicit formula may actually reduce the accuracy, so there might also be cases where the MWH method is less accurate than the SWH method. The MWH method calculates the value of an option with high accuracy, provided that a sufficient number of terms are included in the truncated products.

Another important consideration for a numerical algorithm, particularly in rapidly evolving fields such as financial mathematics, is the computation speed. Our MWH method iteratively calculates the option value using the fast Fourier transform (FFT), which is an efficient algorithm that is highly optimised in most coding languages. Ac-

cordingly, similar to the SWH method that also utilises the FFT, the MWH method is highly efficient and can be implemented with high computation speed. Further, given the calculated Wiener-Hopf factors, the MWH and SWH methods are almost identical, except that the SWH has an additional set of FFTs to implement. Accordingly, the iterative evaluation in the MWH method has higher efficiency than that in the SWH method. However, as with the accuracy, this does not imply that the MWH method is always faster than the SWH method, since calculating the infinite product for the Wiener-Hopf factors in the MWH method may be slower than the approximation in the SWH method, particularly if the truncation length is large. Consequently, it is important that an appropriate truncation length is chosen. If too few terms are used, the accuracy is reduced, while too many terms decreases the speed of the algorithm. Unfortunately, there is no generalised method of choosing this parameter and it currently relies on trial and error. Additionally, the algorithm could be improved if an acceleration technique was derived for the product in the Wiener-Hopf factors. While these goals were not achieved in this investigation, the speed of the MWH still significantly improves on naive methods.

Finally, while we have seen that the MWH method may not have improved speed or accuracy compared to the SWH method, we can apply our method to certain classes of Lévy processes that are unattainable with the SWH method. Recall that the SWH method requires the underlying Lévy process to be a pure jump process with finite jump activity, but the MWH method does not have this restriction. In Section 3.2, we saw that we want to have access to algorithms that are driven by either finite or infinite variation processes since there are financial securities that are suited to each of these cases. Accordingly, this versatility of the MWH method is a significant advantage, while still maintaining a similar speed and accuracy to the SWH method.

Our new MWH method estimates the value of a barrier option where the underlying financial security is modelled according to a meromorphic Lévy process. We have seen that the class of meromorphic Lévy processes provides a wide range of properties and would be able to effectively model many securities. While there are some cases where the MWH method may be slower or less accurate than the SWH method, one of the more recent algorithms in the literature, there are others where the MWH method is more effective, and also some situations where the MWH method can be applied while the SWH method cannot. Overall, the MWH method is an effective algorithm that can be applied when the underlying security is modelled by a meromorphic Lévy process, with high speed and accuracy, and is particularly useful for processes of infinite variation, where some other techniques cannot be applied.

Chapter 7

Numerical experiments

To complete our investigation, the new meromorphic-driven option pricing algorithm is demonstrated through numerical experiments. The evaluation time and accuracy are examined for a range of parameter values that represent interesting cases. The target option value is also calculated using a Monte Carlo framework, where simulations are produced with the Wiener-Hopf Monte Carlo (WHMC) techniques from Sections 4.5 and 6.4. Additionally, to demonstrate the improvement in speed, the results are contrasted to the output from the Simple Wiener-Hopf (SWH) method in Chapter 4.4.

Since the strike price, barrier, and maturity do not influence the comparison of our methods, these parameters are constant for all of our experiments. Consider a knock-out put barrier option with strike $K = 100$, barrier $H = 90$, maturity $T = 1$ year, and with spot price, S_0, taking values 91, 101, 111, 121, and 131. This is the same situation as was investigated using the KoBoL process in Section 4.4.1 and in [32, 33]. While we do not expect identical results since we are using a different underlying process, these values represent situations that are likely to appear in real-world option contracts.

All of our numerical experiments use $N_{pw} = 800$ iterations of the Post-Widder formula, for both the MWH and SWH methods. Increasing this parameter by a factor of 10 resulted in a change to the estimated option value of less than 0.0001 in all of our experiments, so we assume that it is sufficient. Monte Carlo estimates are achieved by using the WHMC technique to produce 5000 realisations of the underlying meromorphic Lévy process, and the resultant option payout is averaged as in standard Monte Carlo applications. Since the simulation of meromorphic Lévy processes is relatively complex, our implementation of the WHMC technique was inefficient and additional simulation was unachievable within a reasonable time frame. Consequently, there is a significant margin of error in these Monte Carlo estimates. However, the estimate does provide an approximate target value for the option price.

First, let the underlying security be modelled by an exponential β-family process, from Section 5.2.1, with parameters $c_1 = c_2 = 100$, $\alpha_1 = \alpha_2 = 5$, $\beta_1 = \beta_2 = 8$, $\lambda_1 = \lambda_2 = 0.5$, and $\sigma = 0$, where both the positive jumps and negative jumps are truncated

to summations of length $N = \widehat{N} = 500$. We consider this truncation to be sufficient for the process to converge since the characteristic exponent and Wiener-Hopf factors both change by less than 0.001 with a tenfold increase in N and $\widehat{N}$. According to the function **calc_mu** that was defined in Section 6.2 and implemented below, we require $\mu = 0.0671$ in this case.

```
sigma = 0
alpha1 = alpha2 = 5
beta1 = beta2 = 8
lambda1 = lambda2 = 0.5
c1 = c2 = 100
r = 0.07231
N = 500

rho = np.array([beta1*(alpha1 + n) for n in range(N)])
rho_hat = np.array([beta2*(alpha2 + n) for n in range(N)])
a = np.array([c1/(beta1*(alpha1 + n)) * sc.binom(n + lambda1 - 1, n)
              for n in range(N)])
a_hat = np.array([c2/(beta2*(alpha2 + n)) * sc.binom(n + lambda2 - 1, n)
                  for n in range(N)])

mu = calc_mu(sigma, a, rho, a_hat, rho_hat, r, dividend = 0)
```

The MWH, SWH, and WHMC methods are applied to estimate the value of this option, using the techniques and code in Sections 6.2, 6.3, and 6.4, respectively. The results of this experiment are shown in Table 7.1. Even with a relatively low number of trials, WHMC required a long time (several days) to produce a result. This was significantly improved by both the SWH method and our new MWH method. Averaging over the separate implementations for the 5 initial spot prices, the SWH and MWH methods had execution times of 3.53 and 5.58 seconds, respectively. While the SWH method was slightly faster in this case, there are other situations where the MWH may be faster and also where the MWH can be applied while the SWH cannot, as discussed in Section 6.5. Thus, the new method is still beneficial. Table 7.1 shows that the three methods produce results that are consistent with each other, particularly the MWH and SWH methods. While the Monte Carlo estimate is slightly different, this is likely due to the relatively low sample size and consequently low accuracy. A 95% confidence interval for each of the Monte Carlo estimates does contain the results from both the MWH and SWH methods, so this experiment suggests that the methods are accurate.

The speed of the MWH and SWH methods both depend on the number of roots that are included, N and $\widehat{N}$. Thus, consider another numerical experiment where the MWH and SWH methods are applied with the same underlying meromorphic Lévy process, but with $N = \widehat{N} = 20$ roots instead of 500. The results of this experiment are shown in Table 7.2. Since the option values are different from those in Table 7.1, the truncation length

of $N = 20$ is not sufficient for the process to converge and we are now considering the corresponding hyper-exponential process, discussed in Section 5.2.2. This is demonstrated in Figure 7.1, where the option price is calculated for a range of values for N, appearing to converge around $N = 200$. Despite being different from the previous experiment, the results given by the MWH and SWH methods are still consistent, within a relatively small margin of error. Furthermore, observe that the speed of the MWH method has improved significantly more than the SWH method. In this case, the MWH method is faster than the SWH method, with computation times of 2.72 and 3.37 seconds, respectively. This experiment confirms that there are situations where the MWH method does improve the speed of previous algorithms, in addition to the other benefits that have been discussed.

Table 7.1: Estimated option values using Monte Carlo simulations (with 95% confidence interval), the MWH method, and the SWH method.

		Monte Carlo	**MWH**	**SWH**
Option price	$\mathbf{S_0 = 91}$	0.539 ± 0.072	0.538	0.541
	$\mathbf{S_0 = 101}$	1.057 ± 0.088	1.066	1.053
	$\mathbf{S_0 = 111}$	0.366 ± 0.058	0.375	0.371
	$\mathbf{S_0 = 121}$	0.074 ± 0.024	0.090	0.086
	$\mathbf{S_0 = 131}$	0.020 ± 0.014	0.017	0.017
Average time (sec)			5.58	3.53

Option parameters: $K = 100$, $H = 90$, $T = 1$, $r = 0.07231$.
Meromorphic model parameters: $c_1 = c_2 = 100$, $\alpha_1 = \alpha_2 = 5$, $\beta_1 = \beta_2 = 8$, $\lambda_1 = \lambda_2 = 0.5$, $\sigma = 0$, $\mu = 0.0671$, $N = \widehat{N} = 500$.

Table 7.2: Estimated option values using the MWH and SWH method.

		MWH	**SWH**
Option price	$\mathbf{S_0 = 91}$	0.658	0.636
	$\mathbf{S_0 = 101}$	1.116	1.110
	$\mathbf{S_0 = 111}$	0.356	0.353
	$\mathbf{S_0 = 121}$	0.077	0.076
	$\mathbf{S_0 = 131}$	0.015	0.015
Average time (sec)		2.72	3.37

Option parameters: $K = 100$, $H = 90$, $T = 1$, $r = 0.07231$.
Meromorphic model parameters: $c_1 = c_2 = 100$, $\alpha_1 = \alpha_2 = 5$, $\beta_1 = \beta_2 = 8$, $\lambda_1 = \lambda_2 = 0.5$, $\sigma = 0$, $\mu = 0.0671$, $N = \widehat{N} = 20$.

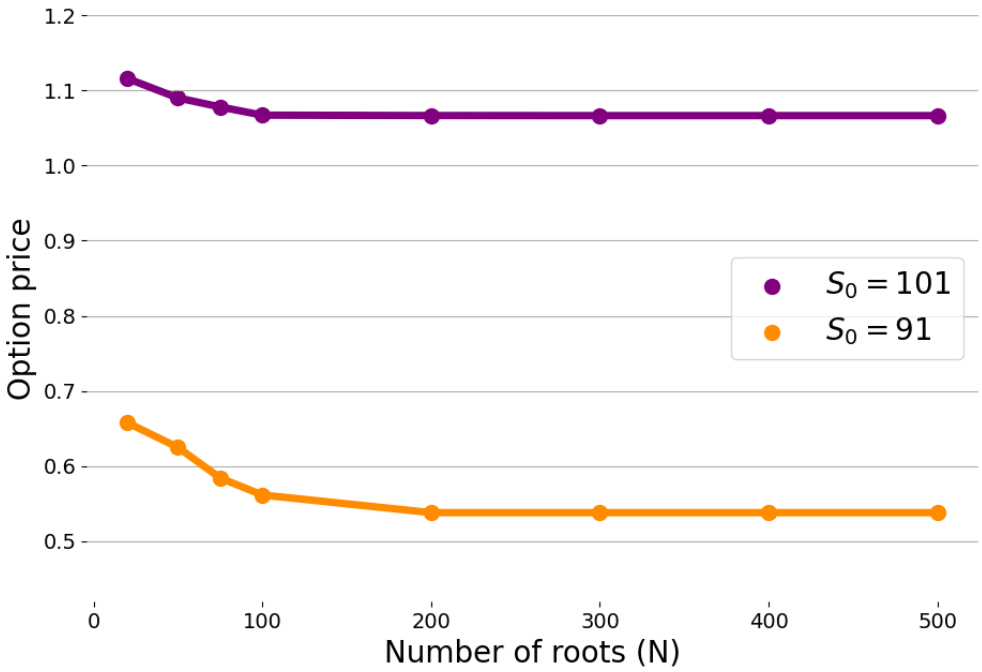


Figure 7.1: Option price calculated for the parameters given in Tables 7.1 and 7.2 using the MWH method with varying number of roots, $N \in [50, 500]$.

Lastly, suppose that we have a different underlying meromorphic Lévy process, with parameters $c_1 = 100$, $c_2 = 80$, $\alpha_1 = \alpha_2 = 4$, $\beta_1 = \beta_2 = 6$, $\lambda_1 = 0.7$, $\lambda_2 = 0.5$, and $\sigma = 0.05$. In this case, we require drift $\mu = 0.0488$. Notice that this process now has a Gaussian component since the diffusivity is non-zero. Thus, the SWH method, which is restricted to non-Gaussian processes, cannot be applied. However, the MWH method can still be used in this case, with results shown in Table 7.3. This experiment demonstrates where our new MWH method may be applied in a situation that could not be solved with the previous SWH method.

Table 7.3: Estimated option values using the MWH method with a non-Gaussian component.

$\mathbf{S_0}$	**91**	**101**	**111**	**121**	**131**	**141**	**151**
Option price	0.020	0.193	0.404	0.545	0.567	0.500	0.394

Option parameters: $K = 100$, $H = 90$, $T = 1$, $r = 0.07231$.
Meromorphic model parameters: $c_1 = 100$, $c_2 = 80$, $\alpha_1 = \alpha_2 = 4$, $\beta_1 = \beta_2 = 6$, $\lambda_1 = 0.7$, $\lambda_2 = 0.5$, $\sigma = 0.05$, $\mu = 0.0488$, $N = \widehat{N} = 500$.

Chapter 8

Conclusion

This research aimed to investigate a relatively new class of stochastic processes called meromorphic Lévy processes and then determine if it may be applied to model financial securities, and hence, estimate the price of barrier options. We were able to derive an option pricing model driven by meromorphic Lévy processes based on a numerical algorithm where the expected value of the option payout is evaluated at each arrival of a Poisson process. The results indicate that our new algorithm, which we call the Meromorphic Wiener-Hopf (MWH) method, has similar speed and accuracy to other recent methods in the literature, while also being able to be applied in some useful additional cases.

Option pricing models driven by stochastic processes have been researched thoroughly in the literature for several decades. Original models involved continuous processes such as Brownian motion, but various issues have led to the use of more generalised Lévy processes. Empirical studies of financial markets in the literature have shown that càdàg processes whose increments have heavy tails are more appropriate. Both bounded and infinite variation processes are useful for different financial markets, so models that can be driven by processes with each of these properties are desirable. We showed that these requirements are all satisfied by the meromorphic class of Lévy processes, which motivated our interest in these processes. We expect the resultant model to more accurately describe some real-world financial markets.

The literature has demonstrated many ways to derive barrier option pricing methods driven by Lévy processes, which can be categorised into three groups: Monte Carlo simulation, backward induction, and integro-differential equations solved using Wiener-Hopf factorisation. Our model was derived using methodology from the latter category because the class of meromorphic Lévy processes has known results for the Wiener-Hopf factorisation. Through this, we were able to derive a new barrier option pricing technique, the MWH method.

Numerical experiments comparing our new MWH method to another recent technique that can use meromorphic Lévy processes, called the Simple Wiener-Hopf (SWH) method, demonstrated that our technique produced the expected results. We also showed that our

MWH method can be applied in situations where the SWH method cannot, so it is a beneficial addition to the literature.

While the MWH method was shown to be successful, there are several areas for future research that could improve, or build upon, this research. In particular, implementation of the semi-explicit Wiener-Hopf factors for meromorphic Lévy processes could be improved, additional subclasses of meromorphic Lévy processes should be investigated, and additional research into fitting meromorphic Lévy processes to given financial securities would be necessary to use this algorithm in practice.

Firstly, an important area of future research would be to improve the implementation of the Wiener-Hopf factors for meromorphic Lévy processes by determining an appropriate truncation length and an acceleration technique for the infinite product. The known expression for the Wiener-Hopf factors is a convergent infinite product that needs to be truncated for implementation. We did not provide an appropriate length for this truncation and instead relied on trial and error to determine when the product became relatively constant. This process was tedious and often does not achieve the optimal truncation length. Further research into the expression for the Wiener-Hopf factors of meromorphic Lévy processes to determine an appropriate truncation length would streamline this process and ensure that the evaluation is optimal. Additionally, computation of the truncated infinite product is slow, meaning that the computation speed of our algorithm was not significantly improved over other methods that approximate the Wiener-Hopf factors. Accordingly, it would be beneficial if future research could derive an acceleration technique for the expressions, similar to that which is available for the Post-Widder formula. This would allow the Wiener-Hopf factors to be calculated much more efficiently and improve the speed of the MWH method. Overall, further research into the expression for the Wiener-Hopf factors of meromorphic Lévy processes would be extremely beneficial for our MWH method, as well as many other applications that rely on the Wiener-Hopf factorisation.

Another area for future research is in categorising additional subclasses of meromorphic Lévy processes. The class of meromorphic Lévy processes is extremely rich and includes processes with a wide range of properties that can be used to define interesting subclasses, such as the β-family, hyper-exponential processes, and θ-family that we analysed. Since this class of Lévy processes was defined relatively recently, there remain many potentially interesting subclasses to be discovered. Future research into subclasses of meromorphic Lévy processes would likely identify processes that may be particularly useful for modelling financial securities or in various other applications.

Finally, techniques to choose meromorphic Lévy processes and fit them to model given financial securities is an important area for future research. While we have shown that the class of meromorphic Lévy processes satisfies the properties that are important for barrier option pricing, we did not discuss how a particular process, and its parameters, would be chosen for a given security. We were more interested in investigating the meromorphic

Lévy processes and their general application in financial markets, not in fitting them to particular securities. However, the ability to choose the right models would significantly enhance the practical use of our algorithm, so this is an important area for future research.

Overall, we successfully derived a new method for calculating the value of barrier options that are modelled according to meromorphic Lévy processes. This demonstrated one of many useful applications of the Wiener-Hopf factorisation and, hence, showed the importance of the class of meromorphic Lévy processes.

Appendix A

Appendix

A.1 Meromorphic function

In the field of complex analysis, a function is meromorphic if it is holomorphic everywhere on its domain, except for a set of isolated points that are poles of the function. This is stated in Definition 46 [18].

Definition 46. *Let $\overline{C} = \mathbb{C} \cup \{\infty\}$ be the Riemann sphere and $D \subset \mathbb{C}$ be open. A map*

$$f \colon D \to \overline{C}$$

is a meromorphic function if:

1. *The set $S(f) = f^{-1}(\{\infty\})$ of points with function value ∞ is discrete in D.*
2. *The restriction, f_0, of f from D to $D \backslash S(f)$ is analytic.*
3. *The points in $S(f)$ are poles of f_0.*

Any meromorphic function, f, may be written as a ratio of two holomorphic functions,

$$f(z) = \frac{g(z)}{h(z)}.$$

If the domain is connected, then the set of meromorphic functions is the field of fractions of the integral domain of holomorphic functions [24]. This suggests that the relationship between the sets of meromorphic functions and holomorphic functions is analogous to the relationship between the rational and integer numbers.

A notable subset of meromorphic functions arises by taking the holomorphic functions g and h to be polynomials. This defines the set of rational functions. The set of poles in this case is finite and corresponds to the zeros of the polynomial h.

We may also have meromorphic functions with a (countably) infinite number of poles. As mentioned by Freitag [18], a typical example of such a meromorphic function is $f(z) = \cot(\pi z) = \frac{\cos(\pi z)}{\sin(piz)}$. This function satisfies Definition 46, where the set of poles $S(f) = \mathbb{Z}$.

A.2 Fast Fourier transform

The Fourier transform of a function, $f\colon \mathbb{R} \to \mathbb{C}$ is

$$\mathcal{F}\{f\}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} f(x)\, dx, \quad \text{for } \xi \in \mathbb{R},$$

and the inverse Fourier transform is

$$\mathcal{F}^{-1}\{\hat{f}\}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} \hat{f}(\xi)\, d\xi, \quad \text{for } x \in \mathbb{R}.$$

The *discrete Fourier transform* is a method to evaluate the Fourier transform when given a set of data from an unknown function. This is efficiently implemented using the *fast Fourier transform*.

Discrete Fourier transform

The discrete Fourier transform (DFT) allows the Fourier transform to be calculated from a set of data instead of a function. Suppose that we have a set of data $[f_0,\, f_1,\, \ldots,\, f_{n-1}]^\top \in \mathbb{R}^n$. The DFT of this data is another vector $[\widehat{f_0},\, \widehat{f_1},\, \ldots,\, \widehat{f_{n-1}}]^\top \in \mathbb{C}^n$, where

$$\widehat{f_k} = \sum_{j=0}^{n-1} f_j e^{-2\pi ijk/n}. \tag{A.1}$$

As with the Fourier transform, we also have an inverse discrete Fourier transform (iDFT), given by

$$f_k = \frac{1}{n} \sum_{j=0}^{n-1} \widehat{f_j} e^{2\pi ijk/n}. \tag{A.2}$$

This is applied instead of the Fourier transform in many numerical implementations, but the DFT is computationally inefficient, with a complexity of $\mathcal{O}(n^2)$.

Fast Fourier Transform

The Fast Fourier transform (FFT) and the inverse fast Fourier transform (iFFT) are algorithms that implement the DFT and iDFT more efficiently. Letting $\omega_n = e^{-2\pi i/n}$, the DFT in (A.1) can be rewritten,

$$\widehat{f}_k = \sum_{j=0}^{n-1} f_j e^{-2\pi i j k/n} = \sum_{j=0}^{n-1} f_j \omega_n^{jk},$$

or, equivalently,

$$\begin{bmatrix} \widehat{f}_0 \\ \widehat{f}_1 \\ \widehat{f}_2 \\ \vdots \\ \widehat{f}_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega_n & \omega_n^2 & \dots & \omega_n^{n-1} \\ 1 & \omega_n^2 & \omega_n^2 & \dots & \omega_n^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{n-1} & \omega_n^{2(n-1)} & \dots & \omega_n^{(n-1)^2} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{bmatrix}.$$

The complexity of this matrix operation can be improved by diagonalisation, as follows. Let

$$F_k = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega_n & \omega_n^2 & \dots & \omega_n^{k-1} \\ 1 & \omega_n^2 & \omega_n^2 & \dots & \omega_n^{2(k-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{k-1} & \omega_n^{2(k-1)} & \dots & \omega_n^{(k-1)^2} \end{bmatrix}, \text{ and } D_k = \begin{bmatrix} 1 & & & & \\ & \omega_n & & 0 & \\ & & \omega_n^2 & & \\ & 0 & & \ddots & \\ & & & & \omega_n^{k-1} \end{bmatrix}.$$

Observe that the DFT is now equivalent to

$$\widehat{f} = F_n f.$$

Then, assuming that $n = 2^m$ for some $m \in \mathbb{Z}_+$, the FFT decomposes the problem to be

$$\widehat{f} = \begin{bmatrix} I_{n/2} & D_{n/2} \\ I_{n/2} & -D_{n/2} \end{bmatrix} \begin{bmatrix} F_{n/2} & 0 \\ 0 & F_{n/2} \end{bmatrix} \begin{bmatrix} f_{\text{even}} \\ f_{\text{odd}} \end{bmatrix},$$

where $f_{\text{even}} = [f_0, f_2, ..., f_{n-2}]$, $f_{\text{odd}} = [f_1, f_3, ..., f_{n-1}]$, and I_n is the $n \times n$ identity matrix. This decomposition is repeated until we have 2×2 matrices, which gives the FFT. The matrices in this procedure can be calculated much more efficiently, giving the FFT a complexity of $\mathcal{O}(n \log(n))$.

Due to the importance of the Fourier transform, the FFT is inbuilt in most coding languages. In Python, the FFT is available in the scipy module as **fft** and **ifft**.

```
from scipy.fft import fft, ifft
```

A.3 Interval bisection method

The interval bisection method is a simple method for calculating the roots of a given continuous function, $f(x)$, for which one knows two points, a and b, with opposite signs. Assume without loss of generality, $f(a) > 0$ and $f(b) < 0$. Then, by the *intermediate value theorem*, there exists a point $x^* \in (a, b)$ that is a root of the function; $f(x^*) = 0$. Let $c = (a + b)/2$ be the midpoint of the interval (a, b). If $f(c) < 0$, we have a root in (a, c). Otherwise, there is a root in (c, b). Thus, this process can be repeated for a new interval with half the width. Iterating this process until an interval of the required width is achieved allows us to determine the root of the function with arbitrary accuracy. We implement the interval bisection method in Python as follows.

```
def bisection(f, a, b, epsilon, max_iter):
    '''Approximate solution of f(x) = 0 by Interval Bisection method.

    Parameters
    ----------
    f : function
        Function for which we are searching for a solution f(x)=0.
    a : number
        Lower endpoint of initial interval.
    b : number
        Upper endpoint of initial interval.
    epsilon : number
        Stopping criteria is abs(f(x)) < epsilon.
    max_iter : integer
        Maximum number of iterations of Newton's method.

    Returns
    -------
    xn : number
        Implement Interval Bisection method.
    '''

    for n in range(0, max_iter):
        c = (a+b)/2
        fxn = f(c)
        if abs(fxn) < epsilon:
            print('Found solution after',n,'iterations.')
            return c
        if np.sign(f(a)) == np.sign(f(c)):
            a = c
        else:
            b = c
    print('Reached maximum iterations.')
    return c
```

A.4 Acceptance-rejection sampling

Acceptance-rejection sampling (or, simply, *rejection sampling*) is a method of generating observations of a random variable, X, with given distribution in $\mathbb{R}^m$ that has a known density, f. Let Y be another random variable that has a density, from which we are able to easily sample (*e.g.* normal, uniform, exponential). Let $M \in (1, \infty)$ be a constant that bounds the likelihood ratio $f(x)/g(x)$, so that, for all x in the support of X,

$$f(x) \leq Mg(x).$$

Thus, we need Y such that some scaled version of the density of Y can dominate the density of X over its entire support. Then, an observation, y, of Y is generated and accepted as a sample of X with probability $f(y)/Mg(y)$. Otherwise, the observation is rejected and the process is repeated until we successfully sample a point. The efficiency of the algorithm depends on the probability of acceptance. If this value is low, many observations are rejected and the algorithm has low efficiency. Thus, the algorithm is optimised if the distribution of Y has a similar shape to that of X and if the value of M is chosen to be the smallest possible bound. The acceptance-rejection sampling algorithm can be stated as follows.

Step 1: Obtain a sample, y, from Y.

Step 2: Sample a uniform random variable $u \sim \text{Unif}(0, 1)$.

Step 3: If $u < \dfrac{f(y)}{Mg(y)}$:

accept y as a sample of X;

Step 4: otherwise:

reject y and return to Step 1.

Bibliography

[1] Abate, J. [1995], 'Numerical Inversion of Laplace Transforms of Probability Distributions', *ORSA Journal on Computing* **7**(1), 36–43.

[2] Bachelier, L. [1900], 'Theory of Speculation', *Annales scientifiques de l'Ecole Normale Supèrieure* **3**(17), 21–86.

[3] Billingsley, P. [1995], *Probability and Measure*, 3rd edn, John Wiley and Sons, Inc.

[4] Bingham, N. and Kiesel, R. [1998], *Risk-Neutral Valuation*, Springer Finance.

[5] Black, F. and Scholes, M. [1973], 'The Pricing of Options and Corporate Liabilities', *Journal of Political Economy* **81**(3), 637–654.

[6] Bogachev, V. [2007], *Measure Theory*, Vol. 1, Springer.

[7] Boyarchenko, M. and Levendorskiĭ, S. [2009], 'Prices and Sensitivities of Barrier and First-Touch Digital Options in Lévy-driven Models', *International Journal of Theoretical and Applied Finance* **12**(8), 1125–1170.

[8] Boyarchenko, S. and Levendorskiĭ, S. [2002], *Non-Gaussian Merton-Black-Scholes Theory*, World Scientific, New Jersey.

[9] Boyarchenko, S. and Levendorskiĭ, S. [1999], 'Option Pricing for Truncated Lévy Processes', *International Journal of Theoretical and Applied Finance* **3**(3), 549–552.

[10] Brummelhuis, R. and Chan, R. T. L. [2014], 'A Radial Basis Function Scheme for Option Pricing in Exponential Lévy Models', *Applied Mathematical Finance* **21**(3), 238–269.

[11] Carr, P. [1998], 'Randomization and the American Put', *Review of Financial Studies* **11**, 105–110.

[12] Chung, K. and Williams, R. [1990], *Introduction to Stochastic Integration*, 2nd edn.

[13] Cont, R. and Voltchkova, E. [2005], 'A Finite Difference Scheme for Option Pricing in Jump Diffusion and Exponential Lévy Models', *SIAM Journal on Numerical Analysis* **43**(4), 1596–1626.

[14] Dhaene, J., Denuit, M., Goovaerts, M., Kaas, R. and Vyncke, D. [2002], 'The Concept of Comonotonicity in Actuarial Science and Finance: Applications', *Insurance: Mathematics and Economics* **31**(2), 133–161.

[15] Drake, P. and Fabozzi, F. [2010], *The Basics of Finance*, The Frank J Fabozzi Series, John Wiley & Sons.

[16] Eberlein, E. [2009], *Jump-Type Lévy Processes*, Springer, Berlin, Heidelberg.

[17] Finetti, B. D. [1929], 'Sulle Funzioni ad incremento aleatorio', *Rend. Acc. Naz. Lincei.* **10**, 163–168.

[18] Freitag, E. and Busam, R. [2009], *Complex Analysis*, Universitext, 2nd edn, Springer Berlin.

[19] Frey, R. [2009], 'Lecture Notes Continuous-Time Finance'. University of Leipzig.

[20] Hamza, K., Jacka, S. and Klebaner, F. [2005], 'The EMM Conditions in a General Model for Interest Rates', *Advances in Applied Probability* **37**(2), 415–434.

[21] Hwang, I. [2021], 'A brief History of the Stock Market'. SoFi Learn.

[22] Itkin, A. [2017], *Pricing Derivatives Under Lévy Models: Modern Finite-Difference and Pseudo-Differential Operators Approach*, 1st edn, Birkhauser.

[23] Jackson, K. R., Jaimungal, S. and Surkov, V. [2008], 'Fourier Space Time-Stepping for Option Pricing with Lévy Models', *Journal of Computational Finance* **12**(2), 1–28.

[24] Judson, T. [2022], *Abstract Algebra: Theory and Applications*, Stephen F. Austin State University.

[25] Kallenberg, O. [1997], *Foundations of Modern Probability*, Springer.

[26] Karatzas, I. and Shreve, S. [1998], *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, 2nd edn, Springer, New York.

[27] Kirkby, J. L. [2017], 'Robust Option Pricing with Characteristic Functions and the B-spline Order of Density Projection', *Journal of Computational Finance* **21**(2), 61–100.

[28] Kou, S. [2002], 'A Jump Diffusion Model for Option Pricing', *Management Science* pp. 1086–1101.

[29] Kudryavtsev, O. [2011], 'An Efficient Numerical Method to Solve a Special Class of Integro-differential Equations Relating to the Lévy Models', *Mathematical Models and Computer Simulations* **3**(6), 706–711.

[30] Kudryavtsev, O. [2016], 'Advantages of the Laplace Transform Approach in Pricing First Touch Digital Options in Lévy-driven Models', *Boletín de la Sociedad Matemática Mexicana* **22**(2), 711–731.

[31] Kudryavtsev, O. E. [2019], 'Approximate Wiener-Hopf Factorization and Monte Carlo Methods for Lévy Processes', *Theory of Probability and its Applications* **64**(2), 186–208.

[32] Kudryavtsev, O. and Levendorskiĭ, S. [2009], 'Fast and Accurate Pricing of Barrier Options Under Lévy Processes', *Finance and Stochastics* **13**, 531–562.

[33] Kudryavtsev, O. and Luzhetskaya, P. [2020], 'The Wiener-Hopf Factorization for Pricing Options Made Easy', *IAENG Engineering Letters* **28**(4), 1310–1317.

[34] Kuznetsov, A. [2010*a*], 'Wiener-Hopf Factorization and Distribution of Extrema for a Family of Lévy Processes', *Annals of Applied Probability* **20**, 1801–1830.

[35] Kuznetsov, A. [2010*b*], 'Wiener-Hopf Factorization for a Family of Lévy Processes Related to Theta Functions', *Journal of Applied Probability* **47**(4), 1023–1033.

[36] Kuznetsov, A., Kyprianou, A. E. and Pardo, J. C. [2012], 'Meromorphic Lévy Processes and Their Fluctuation Identities', *Annals of Applied Probability* **22**(3), 1101–1135.

[37] Kuznetsov, A., Kyprianou, A. E., Pardo, J. P. and van Schaik, K. [2011], 'A Wiener-Hopf Monte Carlo Simulation Technique for Lévy Processes', *Annals of Applied Probability* **21**(6), 2171–2190.

[38] Kyprianou, A. E. [2014], *Fluctuations of Lévy Processes With Applications*, Universitext, 2nd edn, Springer.

[39] Lévy, P. [1937], 'Théorie de l'addition des variables aléatoires', *Bulletin de la Société Mathématique de France* **67**(1), 1–41.

[40] Lukacs, E. [1970], *Characteristic Functions*, 2nd edn, Hafner Publishing Co., New York.

[41] Mandelbrot, B. [1963], 'New Methods in Statistical Economics', *Journal of Political Economy* **71**(5), 421–440.

[42] Merton, R. [1973], 'Theory of Rational Option Pricing', *The Bell Journal of Economics and Management Science* **4**(1), 141–183.

[43] Najafabadi, A. T. P. [2018], 'A Mixture-Exponential Approximation to Finite and Infinite-time Ruin Probabilities of Compound Poisson Processes', *IAENG International Journal of Applied Mathematics* **48**(1), 105–110.

[44] National Museum of Australia [2022], 'Defining Moments: First Stock Exchanges', www.nma.gov.au/defining-moments/resources/first-stock-exchange. Accessed on 22/09/2022.

[45] Phelan, C. E., Marazzina, D., Fusai, G. and Germano, G. [2018], 'Fluctuation Identities with Continuous Monitoring and Their Application to Price Barrier Options', *European Journal of Operational Research* **271**(1), 210–223.

[46] Pinsky, M. A. and Karlin, S. [2011], *An Introduction to Stochastic Modeling*, 4$^{\text{th}}$ edn, Academic Press.

[47] Platen, E. and Liberati, N. [2010], *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, Vol. 64 of *Stochastic Modelling and Applied Probability*, Springer Berlin, Heidelberg.

[48] Privault, N. [2022], *Notes on Stochastic Finance*, Nanyang Technological University.

[49] Rogers, L. [1983], 'Wiener-Hopf Factorisation of Diffusions and Lévy Processes', *Proceedings of the London Mathematical Society* **47**(3), 177–191.

[50] Ruijter, M. J. and Oosterlee, C. W. [2016], 'Numerical Fourier Method and Second-order Taylor Scheme for Backward SDES in Finance', *Applied Numerical Mathematics* **103**(C), 1–26.

[51] Sato, K. [2000], *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

[52] Staum, J. [2008], 'Incomplete Markets', *Handbooks in Operations Research and Management Science* **15**(7), 511–563.

[53] Valdez, S. and Molyneux, P. [2016], *An Introduction to Global Financial Markets*, 8$^{\text{th}}$ edn, Palgrave.

[54] Widder, D. [1947], *Advanced Calculus*, Prentice-Hall Mathematics Series, Prentice-Hall, Inc.

[55] Williams, R. J. [2006], *Introduction to the Mathematics of Finance*, Vol. 72 of *Graduate Studies in Mathematics*, American Mathematical Society.

[56] Yang, Y., Ma, J. and Liang, Y. [2018], 'The Research on the Calculation of Barrier Options under Stochastic Volatility Models Based on the Exact Simulation', *IAENG International Journal of Applied Mathematics* **48**(3), 349–361.